PRELIMINARY AND INCOMPLETE*

A Unique Costly Contemplation Representation

Haluk Ergin† Todd Sarver‡

November 3, 2005

Abstract

We extend the costly contemplation model to preferences over sets of lotteries, assuming that the state-dependent utilities are von Neumann-Morgenstern. The contemplation costs are uniquely pinned down in a reduced form representation, where the decision-maker selects a subjective measure over expected utility functions instead of a subjective signal over a subjective state space. We show that in this richer setup, costly contemplation is characterized by Aversion to Contingent Planning, Indifference to Randomization, Translation Invariance, and Strong Continuity.

*Please do not circulate without the permission of the authors.

†Correspondence address: MIT, Department of Economics, E52-383A, 50 Memorial Drive, Cambridge, MA 02142. Email: hergin@mit.edu.

‡Correspondence address: Boston University, Department of Economics, 270 Bay State Road Boston, MA 02215. Email: sarvertd@bu.edu.

1 Introduction

1.1 Brief Overview

We will take a Decision Maker (DM) to one of two restaurants. The first one is a seafood restaurant that serves a tuna (t) and a salmon (s) dish, which we denote by $A = \{t, s\}$. The second one is a steak restaurant that serves a filet mignon (f) and a ribeye (r) dish, which we denote by $B = \{f, r\}$. We will flip a coin to determine which restaurant to go to. If it comes up heads then we will buy the DM the meal of her choice in A, if it comes up tails then we will buy her the meal of her choice in B.

We consider presenting the DM one of the two following decision problems:

Decision Problem 1
We ask the DM to make a complete contingent plan listing what she would choose conditional on each outcome of the coin flip.

Decision Problem 2
We first flip the coin and let the DM know its outcome. She then selects the dish of her choice from the restaurant determined by the coin flip.

Decision problem 1 corresponds to a choice out of $A \times B = \{(t, f), (t, r), (s, f), (s, r)\}$, where for instance (s, f) is the plan where the DM indicates that she will have the salmon dish from the seafood restaurant if the coin comes up heads and she will have the filet mignon from the steak restaurant if the coin comes up tails. Note that each choice of a contingent plan eventually yields a lottery over meals. For example if the DM chooses (s, f) then she will face the lottery $\frac{1}{2}s + \frac{1}{2}f$ that yields a salmon and a filet mignon dish, each with one-half probability. Hence decision problem 1 is identical to a choice out of the set of lotteries $\left\{\frac{1}{2}t + \frac{1}{2}f, \frac{1}{2}t + \frac{1}{2}r, \frac{1}{2}s + \frac{1}{2}f, \frac{1}{2}s + \frac{1}{2}r\right\}$.

It is conceivable that the DM prefers facing the second decision problem rather than the first one. In this case we say that her preferences (over decision problems) exhibit *Aversion to Contingent Planning (ACP)*. One explanation of ACP is that the DM finds it psychologically costly to figure out her tastes over meals. Because of this cost, she would rather not contemplate on an inconsequential decision: In our restaurant example, she would rather not contemplate about her choice out of A, were she to know that the coin came up tails and her actual choice set is B. In particular she prefers to learn which choice set (A or B) is the relevant one, before contemplating on her choice.

Our main results are a representation and a uniqueness theorem for preferences over sets of lotteries. We interpret that the preference arises from a choice situation where, initially the DM chooses from among sets of lotteries (menus, options sets, or decision problems) and subsequently chooses a lottery from that set. The only primitive of the model is the preference over sets of lotteries which corresponds to the DM's choice behavior in the first period, we do not explicitly model the second period choice out of the sets. The key axiom in our analysis is ACP and our representation is a reduced form of the costly contemplation representation introduced in Ergin (2003). We begin by a brief overview of Ergin's and our results before we present them more formally in the next section.

The primitive of Ergin's model is a preference over sets of alternatives. He shows that if this preference is monotone, in the sense that each option set is weakly preferred to its subsets, then the DM behaves as if she optimally contemplates her mood before making a choice out of her option set. Ergin models contemplation as a subjective information acquisition problem, where the DM optimally acquires a costly subjective signal over a subjective state space. He interprets these subjective signals as contemplation strategies. The subjective state space, signals, and costs are all parameters of the representation but not a part of the model. However, as we illustrate in the next section these parameters are hardly pinned down from the preference over sets.

We extend the costly contemplation model to preferences over lotteries assuming that the state dependent utilities are von Neumann-Morgenstern. We model a contemplation strategy as a subjective measure over expected utility functions instead of a subjective signal over a subjective state space. The extension of the domain of preferences and the reduction of the parameters of the representation make it possible to uniquely identify contemplation costs from the preference. We show that in this extended model, ACP, indifference to randomization, along with continuity and translation invariance properties characterize costly contemplation. We also prove that the measures in our representation are positive if and only if the preference is monotone.

1.2 Background and detailed results

The costly contemplation representation in Ergin (2003) is the following:

$$V(A) = \max_{\pi \in \Pi} \left[\sum_{E \in \pi} \max_{z \in A} \sum_{\omega \in E} U(z, \omega) - c(\pi) \right] \tag{1}$$

The interpretation of the above formula is as follows. The DM has a finite subjective state space Ω representing her tastes over alternatives. She does not know the realization of the state $\omega \in \Omega$ but has a uniform prior on Ω. Her tastes over alternatives in Z are represented by a state dependent utility function and U: $Z \times \Omega \to \mathbb{R}$. Before making a choice out of a set $A \subset Z$, the DM may engage in contemplation. A *contemplation strat-*

egy is modelled as a signal about the subjective state which corresponds to a partition π of Ω. If the DM carries out the contemplation strategy π, she incurs a psychological cost of contemplation $c(\pi)$, learns which event of the partition π the actual state lies in, and picks an alternative that yields the highest expected utility conditional on each event $E \in \pi$. The set of partitions of Ω is denoted by Π. Faced with the option set A, the DM chooses an optimal level of contemplation by maximizing the value minus the cost of contemplation. This yields $V(A)$ in (1) as the ex-ante value of the option set A.

The appeal of the above formula (1) comes from its similarity to an optimal information acquisition formula. It expresses the costly contemplation problem as a problem with which we are more familiar as economists. The difference from a standard information acquisition problem is that, in the costly contemplation formula the parameters (Ω, U, c) are not directly observable but need to be derived from the DM's preference. Ergin shows that a preference $\succsim$ over sets of alternatives is monotone ($A \subset B \Rightarrow B \succsim A$) if and only if there exist parameters (Ω, U, c) such that $\succsim$ is represented by the ex-ante utility function V in (1). Unfortunately, these parameters are not pinned down from the preference $\succsim$. To illustrate this, first consider the following example:

Example 1 Let $Z = \{z_1, z_2, z_3\}$. All of the following cost functions lead to the same preference over sets of alternatives:

U	ω_1	ω_2	ω_3
z_1	5	0	0
z_2	0	5	0
z_3	0	0	2

	c	c'
$\{\{\omega_1, \omega_2, \omega_3\}\}$	0	0
$\{\{\omega_i\}, \{\omega_j, \omega_k\}\}$	3	4
$\{\{\omega_1\}, \{\omega_2\}, \{\omega_3\}\}$	4	5

given by $\{z_1, z_2, z_3\} \succ \{z_1, z_2\} \succ \{z_1, z_3\} \sim \{z_2, z_3\} \sim \{z_1\} \sim \{z_2\} \succ \{z_3\}$.

The above example should not come as a big surprise. Because of the finiteness of Ergin's model, there is only a finite number preferences over sets, hence the preference data is not enough to pin down one among a continuum possibility of costs. This suggests that the non-uniqueness problem might be resolved by increasing the domain of the preferences to sets of lotteries. Let $\Delta(Z)$ stand for the set of lotteries over Z, and let A now denote a compact set of lotteries. The natural generalization of the costly contemplation formula in (1) to sets of lotteries is:

$$V(A) = \max_{\pi \in \Pi} \left[\sum_{E \in \pi} \max_{p \in A} \sum_{\omega \in E} U(p, \omega) - c(\pi) \right] \tag{2}$$

where the state-dependent utilities $U(\cdot, \omega) \colon \Delta(Z) \to \mathbb{R}$ are assumed to be von Neuman-Morgenstern to guarantee additional structure in the extended model. Assuming that the preference over lotteries has a representation as in (2), it is indeed possible to distinguish between the alternative cost functions in Example 1 from the preference. For instance $V(\{z_1, \frac{1}{2}z_2 + \frac{1}{2}z_3\}) = 5.5 > 5 = V(\{z_1\})$ when the cost function is c, whereas $V'(\{z_1, \frac{1}{2}z_2 + \frac{1}{2}z_3\}) = 5 = V'(\{z_1\})$ when the cost function is c'. However, as we show next, extending the model to sets of lotteries is still not enough to guarantee uniqueness of the parameters (Ω, U, c).

Example 2 Let $Z = \{z_1, z_2\}$ and consider the following two specifications of state spaces $\omega = \{\omega_1, \omega_2, \omega_3\}$, $\hat{\Omega} = \{\hat{\omega}_1, \hat{\omega}_2, \hat{\omega}_3\}$, and the corresponding partition costs:

Ω:

U	ω_1	ω_2	ω_3
z_1	1	-2	2
z_2	-1	2	-2

$c(\{\{\omega_1, \omega_2, \omega_3\}\})$	0
$c(\{\{\omega_1, \omega_2\}, \{\omega_3\}\})$	1
$c(\{\{\omega_1\}, \{\omega_2\}, \{\omega_3\}\})$	2

$\hat{\Omega}$:

$\hat{U}$	$\hat{\omega}_1$	$\hat{\omega}_2$	$\hat{\omega}_3$
z_1	-1	-1	3
z_2	1	1	-3

$\hat{c}(\{\{\hat{\omega}_1, \hat{\omega}_2, \hat{\omega}_3\}\})$	0
$\hat{c}(\{\{\hat{\omega}_1\}, \{\hat{\omega}_2, \hat{\omega}_3\}\})$	1
$\hat{c}(\{\{\hat{\omega}_1\}, \{\hat{\omega}_2\}, \{\hat{\omega}_3\}\})$	2

In each of the two cases above, assume that the cost of all other partitions is 2. Then for any compact set of lotteries A, the parameters (Ω, U, c) and $(\hat{\Omega}, \hat{U}, \hat{c})$ yield the same ex-ante utility value $V(A)$ in (2).

Let U_E denote the expected utility function conditional on an event E, defined by $U_E(p) = \sum_{\omega \in E} U(p, \omega)$. Each partition π induces a collection of such conditional expected utility functions $(U_E)_{E \in \pi}$. When the DM undertakes the contemplation strategy π, she perceives that as a result of her contemplation, she will end up with an ex-post utility function in $(U_E)_{E \in \pi}$. Moreover, it is enough for her to know the cost $c(\pi)$ and the ex-post utility functions $(U_E)_{E \in \pi}$ associated with each contemplation strategy π, to evaluate the ex-ante value $V(A)$ in (2). In particular, it is not possible to behaviorally distinguish between two sets of parameters (Ω, U, c) and $(\hat{\Omega}, \hat{U}, \hat{c})$ that induce the same collections of ex-post conditional utility functions at the same costs.

To be concrete, consider the following table which lists the ex-post utility functions corresponding to the three partitions of Ω in Example 2:

π	$(U_E)_{E \in \pi}$
$\{\{\omega_1, \omega_2, \omega_3\}\}$	$(1, -1)$
$\{\{\omega_1\}, \{\omega_2, \omega_3\}\}$	$(1, -1)$, $(0, 0)$
$\{\{\omega_1, \omega_3\}, \{\omega_2\}\}$	$(3, -3)$, $(-2, 2)$
$\{\{\omega_1, \omega_2\}, \{\omega_3\}\}$	$(-1, 1)$, $(2, -2)$
$\{\{\omega_1\}, \{\omega_2\}, \{\omega_3\}\}$	$(1, -1)$, $(-2, 2)$, $(2, -2)$

where we associate an expected utility function $u\colon \Delta(Z) \to \mathbb{R}$ with the vector $(u(z_1), u(z_2))$.

Let $\mathcal{U}$ denote the non-constant expected utility functions on $\Delta(Z)$, normalized up to positive affine transformations. In Example 2, we can take $\mathcal{U} = \{(1, -1), (-1, 1)\}$.[1] Each ex-post utility function U_E must be a non-negative affine transformation of some $u \in \mathcal{U}$. In particular, each partition π induces a measure μ_π on $\mathcal{U}$, where the weight $\mu_\pi(u)$ of $u \in \mathcal{U}$ is the sum of the non-negative affine transformation coefficients given by:

$$\{\alpha_E \,|\, U_E = \alpha_E u + \beta_E,\ \alpha_E \geq 0, \text{ and } E \in \pi\}.$$

The measures induced by the above partitions are given by:

π	$\mu_\pi(1,-1)$	$\mu_\pi(-1,1)$
$\{\{\omega_1, \omega_2, \omega_3\}\}$	1	0
$\{\{\omega_1\}, \{\omega_2, \omega_3\}\}$	1	0
$\{\{\omega_1, \omega_2\}, \{\omega_3\}\}$	2	1
$\{\{\omega_1, \omega_3\}, \{\omega_2\}\}$	3	2
$\{\{\omega_1\}, \{\omega_2\}, \{\omega_3\}\}$	3	2

We can use these induced measures to rewrite (2) in the alternative form:

$$V(A) = \max_{\mu \in \mathcal{M}} \left[\sum_{u \in \mathcal{U}} \mu(u) \max_{p \in A} u(p) - c(\mu) \right] \tag{3}$$

where the set of partitions Π are replaced by the set of measures $\mathcal{M} = \{\mu_\pi | \pi \in \Pi\}$ over $\mathcal{U}$ and the cost of a measure is defined by $c(\mu) = \min\{c(\pi) \,|\, \mu = \mu_\pi\}$. In this formulation, a contemplation strategy is expressed as measure over expected utility functions instead of a subjective signal π over the subjective state space Ω.

Note that the integral of u with respect to an induced measure gives the ex-ante utility function U_Ω. In the current example $U_\Omega = (1, -1)$. Therefore the measures in $\mathcal{M}$ also satisfy the following *consistency condition*:

$$\forall \mu, \nu \in \mathcal{M}: \qquad \sum_{u \in \mathcal{U}} \mu(u)\, u = \sum_{u \in \mathcal{U}} \nu(u)\, u. \tag{4}$$

Even though a DM's realized tastes *ex-post* contemplation can be very different from her tastes *ex-ante* contemplation, the condition above requires that the contemplation process should not affect the DM's tendencies on the average.

[1] We chose this normalization for the readability of the example, the normalization that we use later for our results is different than the one above.

2 Axioms and Representation

2.1 Axioms

Let Z be a finite set of alternatives and let $\triangle(Z)$ denote the set of all probability distributions on Z. Let $\mathcal{A}$ denote the set of all closed subsets of $\triangle(Z)$ endowed with the Hausdorff metric d_h.[2] Elements of $\mathcal{A}$ are called menus or option sets. The primitive of our model is a binary relation $\succsim$ on $\mathcal{A}$, representing the DM's preferences over menus. We maintain the interpretation that, after committing to a particular menu A, the DM chooses a lottery out of A in an unmodelled second stage. For any $A, B \in \mathcal{A}$ and $\lambda \in [0,1]$, define $\lambda A + (1-\lambda)B \equiv \{\lambda p + (1-\lambda)q : p \in A \text{ and } q \in B\}$. Let $co(A)$ denote the convex hull of the set A.

Axiom 1 (Weak Order): $\succsim$ *is complete and transitive.*

Axiom 2 (Strong Continuity):

1. (Continuity): *For all $A \in \mathcal{A}$, the sets $\{B \in \mathcal{A} : B \succsim A\}$ and $\{B \in \mathcal{A} : B \precsim A\}$ are closed.*
2. (Properness): *There exists $\theta \in \Theta$ such that for all $A, B \in \mathcal{A}$ and $\epsilon > 0$, if $d_h(A,B) < \epsilon$ and $A + \epsilon\theta \in \mathcal{A}$, then $A + \epsilon\theta \succ B$.*

Axioms 1 and 2.1 are entirely standard. Axiom 2.2 states that there exists some $\theta \in \Theta$ such that adding any positive multiple ϵ of θ to A improves this menu, and if any other menu B is chosen "close" enough to A, then $A + \epsilon\theta$ is preferred to B. The role of the strong continuity axiom will essentially be to obtain Lipschitz continuity of our representation in much the same way the continuity axiom is used to obtain continuity. The additional condition imposed in the strong continuity axiom is very similar to the properness condition proposed by Mas-Colell (1986).

The next axiom is introduced in Dekel, Lipman, and Rustichini (2001).

Axiom 3 (Indifference to Randomization (IR)): *For every $A \in \mathcal{A}$, $A \sim co(A)$.*

IR is justified if the DM choosing from the menu A can also randomly select an item from the menu, for example, by flipping a coin. In that case, the menus A and $co(A)$ offer the same set of options, hence they are identical from the perspective of the DM. The next axiom captures an important aspect of our model of costly contemplation.

[2] If we let d denote the Euclidean metric on $\triangle(Z)$, then the Hausdorff metric is defined by

$$d_h(A,B) = \max\left\{\max_{p \in A}\min_{q \in B} d(p,q), \max_{q \in B}\min_{p \in A} d(p,q)\right\}$$

For a full discussion of the Hausdorff metric topology, see Section 3.15 of Aliprantis and Border (1999).

Axiom 4 (Aversion to Contingent Planning (ACP)): *For any $A, B \in \mathcal{A}$, if $A \sim B$ and $\lambda \in (0,1)$, then $A \succsim \lambda A + (1-\lambda)B$.*

Any lottery $p \in \lambda A + (1-\lambda)B$ can be expressed as a mixture $\lambda p_1 + (1-\lambda)p_2$ for some lotteries $p_1 \in A$ and $p_2 \in B$. Hence a choice out of $\lambda A + (1-\lambda)B$ entails choosing from *both* menus. Making a choice out of the menu $\lambda A + (1-\lambda)B$ essentially corresponds to making a contingent plan. The DM's references satisfy ACP if she prefers to be told whether she will be choosing from the menu A or the menu B before actually making her choice. If contemplation is costly, then the individual would prefer not to make such a contingent choice. For example, suppose A is a singleton, so that it is optimal not to contemplate when faced with A, and suppose that it is optimal to contemplate when faced with B. Then, when the individual is faced with the menu $\lambda A + (1-\lambda)B$, she is forced to decide on a level of contemplation before realizing whether she will be choosing from A or B. She would likely choose some intermediate level of contemplation, thus engaging in a level of contemplation that is unnecessarily large (and costly) when the realized menu is A and too small when the realized menu is B. Clearly, she would prefer to be told whether she will be choosing from A or B before deciding on a level of contemplation so that she could contemplate only when necessary.

Before introducing the remaining axioms, we define the set of *singleton translations* to be

$$\Theta \equiv \{\theta \in \mathbb{R}^Z : \sum_{z \in Z} \theta_z = 0\}. \tag{5}$$

For $A \in \mathcal{A}$ and $\theta \in \Theta$, define $A + \theta \equiv \{q + \theta : q \in A\}$. Intuitively, adding θ to A in this sense simply "shifts" A. Also, note that for any $p, q \in \triangle(Z)$, we have $p - q \in \Theta$.

The contemplation decision of an individual will naturally depend on the menu in question. For example, assuming contemplation is costly, an individual will choose the lowest level of contemplation when faced with a singleton menu. As the individual is faced with "larger" menus, we would expect her to choose a higher level of contemplation. However, if a menu is simply shifted by adding some $\theta \in \Theta$ to every item in the menu, this should not alter the individual's desired level of contemplation. Thus preferences will exhibit some form of translation invariance, which we formalize as follows:

Axiom 5 (Translation Invariance (TI)): *For any $A, B \in \mathcal{A}$ and $\theta \in \Theta$ such that $A + \theta, B + \theta \in \mathcal{A}$,*

1. *If $A \succsim B$, then $A + \theta \succsim B + \theta$.*
2. *If $A + \theta \succsim A$, then $B + \theta \succsim B$.*

Finally, we will also consider the monotonicity axiom of Kreps (1979) in conjunction with our other axioms to obtain a refinement of our representation.

Axiom 6 (Monotonicity (MON)): *If $A \subset B$, then $B \succsim A$.*

2.2 Representation

The elements of our representation are a state space S, a Borel measurable state-dependent utility function $U : \triangle(Z) \times S \to \mathbb{R}$, a set of finite signed Borel measures $\mathcal{M}$ on S, and a cost function $c : \mathcal{M} \to \mathbb{R}$. Note that any $\mu \in \mathcal{M}$ is a signed measure, so it can take both positive and negative values. We consider a representation $V : \mathcal{A} \to \mathbb{R}$ defined by

$$V(A) = \max_{\mu \in \mathcal{M}} \left[\int_S \max_{q \in A} U(q,s)\mu(ds) - c(\mu) \right]. \tag{6}$$

We want to think of the different measures as representing different levels of information or contemplation. However, the representation defined in Equation (6) is too general to always fit into this interpretation. We therefore impose the following restriction on the measures in our representation:

Definition 1 *Given $(S, \mathcal{M}, U, c)$, the set of measures $\mathcal{M}$ is said to be* consistent *if for each $\mu, \mu' \in \mathcal{M}$ and $q \in \triangle(Z)$,*

$$\int_S U(q,s)\mu(ds) = \int_S U(q,s)\mu'(ds).$$

It easy to see that this restriction is necessary for our contemplation interpretation. Suppose μ and μ' represent different levels of contemplation, and suppose $q \in \triangle(Z)$. Since the individual has only one choice when faced with the singleton menu $\{q\}$, she cannot change her choice based on her information. Therefore, the only effect of the individual's contemplation decision on her utility is the effect it has on her contemplation cost c. Thus the first term in the brackets in Equation (6) must be the same for both μ and μ', which implies the set of measures $\mathcal{M}$ must be consistent.

We also restrict attention to representations for which each $U(\cdot, s)$ is an expected-utility function. That is, for each $s \in S$, there exists $u \in \mathbb{R}^Z$ such that for all $q \in \triangle(Z)$,

$$U(q,s) = q \cdot u.$$

Define the set of *normalized (non-constant) expected-utility functions* on $\triangle(Z)$ to be $\mathcal{U} = \{u \in \mathbb{R}^Z : \sum_z u_z = 0, \sum_z u_z^2 = 1\}$. Since each $U(\cdot, s)$ is assumed to be an affine function, for each s there exists $u \in \mathcal{U}$ and constants $\alpha, \beta \in \mathbb{R}$ with $\alpha \geq 0$ such that for all $q \in \triangle(Z)$,

$$U(q,s) = \alpha\,(u \cdot q) + \beta.$$

Therefore, by appropriate normalization of the measures and the cost function, we can take $S = \mathcal{U}$ and $U(q, u) = u \cdot q$. That is, it is without loss of generality to assume that the set of states is actually the set of expected-utility functions. While our representation theorems do not require such a restriction, our uniqueness results will be easier to understand if we write the representation in this canonical form. Intuitively, if we do not normalize our expected-utility functions in some way, then there will be an extra "degree of freedom" that must be accounted for in the uniqueness results. Given this normalization, the elements of our representation are simply a pair $(\mathcal{M}, c)$, and the representation in Equation (6) can be written as

$$V(A) = \max_{\mu \in \mathcal{M}} \left[\int_{\mathcal{U}} \max_{q \in A} \big(q \cdot u\big) \mu(du) - c(\mu) \right]. \tag{7}$$

Another important issue related to the assumption that $S = \mathcal{U}$ is whether or not we are imposing a state space on the representation. We argue that the critical restriction here is that each $U(\cdot, s)$ is an expected-utility function, not that $S = \mathcal{U}$. The set $\mathcal{U}$ makes available to the individual all possible expected-utility preferences and therefore does not restrict the state space. It is true that in some cases the set $\mathcal{U}$ may contain "more" states than are necessary to represent an individual's preferences. However, the support of the measures in the representation will reveal which of the expected-utility functions in $\mathcal{U}$ are necessary. Thus, instead of considering minimality of the state space, it is sufficient to consider minimality of the set of measures in the representation, which leads us to our next definition:[3]

Definition 2 *Given a compact set of measures $\mathcal{M}$ and a cost function c, suppose V defined as in Equation (7) represents $\succsim$. The set $\mathcal{M}$ is said to be* minimal *if for any compact proper subset $\mathcal{M}'$ of $\mathcal{M}$, the function V' obtained by replacing $\mathcal{M}$ with $\mathcal{M}'$ in Equation (7) no longer represents $\succsim$.*

We are now ready to formally define our representation:

Definition 3 *A* Reduced Form Costly Contemplation (RFCC) representation *is a compact set of finite signed Borel measures $\mathcal{M}$ and a lower semi-continuous function $c : \mathcal{M} \to \mathbb{R}$ such that*

1. *$V : \mathcal{A} \to \mathbb{R}$ defined by Equation (7) represents $\succsim$.*
2. *$\mathcal{M}$ is both consistent and minimal.*

[3]Note that we endow the set of all finite signed Borel measures on $\mathcal{U}$ with the weak* topology, that is, the topology where a net $\{\mu_d\}_{d \in D}$ converges to μ if and only if $\int_{\mathcal{U}} f\, \mu_d(du) \to \int_{\mathcal{U}} f\, \mu(du)$ for every continuous function $f : \mathcal{U} \to \mathbb{R}$.

3. There exist $p, q \in \triangle(Z)$ such that $V(\{p\}) > V(\{q\})$.

The first two requirements of this definition have been explained above. The third condition is simply a technical requirement relating to the strong continuity axiom. If we take θ as in the definition of strong continuity, then taking any $q \in \triangle(Z)$ and $\epsilon > 0$ such that $q + \epsilon\theta \in \triangle(Z)$ implies $\{q + \epsilon\theta\} \succ \{q\}$, which gives rise to this third condition.

3 Main Results

Theorem 1 *A. The preference $\succsim$ has a RFCC representation if and only if it satisfies weak order, strong continuity, IR, ACP, and TI.*

B. The preference $\succsim$ has a RFCC representation in which each $\mu \in \mathcal{M}$ is positive if and only if it satisfies weak order, strong continuity, MON, ACP, and TI.

Theorem 2 *If $(\mathcal{M}, c)$ and $(\mathcal{M}', c')$ are two RFCC representations for $\succsim$, then there exist constants $\alpha, \beta \in \mathbb{R}$ with $\alpha > 0$ such that $\mathcal{M}' = \alpha\mathcal{M}$ and for all $\mu \in \mathcal{M}$, $c'(\alpha\mu) = \alpha c(\mu) + \beta$.*

Appendix[4]

A Variation of the Mazur Density Theorem

Suppose X is a real Banach space. Before introducing the Mazur density theorem and our variation of the theorem, we need to define the subdifferential of a function:

Definition 4 *Suppose $C \subset X$ is convex and $f : C \to \mathbb{R}$ is a convex function. For $x \in C$, the* subdifferential *of f at x is defined to be*

$$\partial f(x) = \{x^* \in X^* : \langle y - x, x^* \rangle \leq f(y) - f(x) \text{ for all } y \in C\}.$$

The subdifferential is important in the approximation of a convex function by affine functions. In fact, it is straightforward to show that $x^* \in \partial f(x)$ if and only if the affine function $h(y) \equiv f(x) + \langle y - x, x^* \rangle$ satisfies $h \leq f$ and $h(x) = f(x)$. It should also be noted that when X is infinite-dimensional it is possible to have $\partial f(x) = \emptyset$ for some $x \in C$, even if f is convex. However, as proved in Phelps (1993, Proposition 1.11), if f is continuous and convex and C is open, then $\partial f(x) \neq \emptyset$ for all $x \in C$.

[4]Technical Disclaimer: We are currently reorganizing the appendix. Proofs of two Lemmas in Appendix B.3 and of minimality in Appendix B.4 are not complete. The proof of uniqueness is also not available in this version.

The Mazur density theorem states that if X is a separable Banach space and $f : C \to \mathbb{R}$ is a continuous convex function defined on a convex open subset C of X, then the set of points x where $\partial f(x)$ is a singleton is a dense G_δ set in C.[5] The notation G_δ indicates that a set is the countable intersection of open sets.

We wish to obtain a variation of this theorem by relaxing the assumption that C has a nonempty interior. However, it can be shown that the conclusion of the theorem does not hold for arbitrary convex sets. We will therefore require that the affine hull of C, defined below, is dense in X.

Definition 5 *The* affine hull *of a set $C \subset X$, denoted $aff(C)$, is defined to be the smallest affine subspace of X that contains C.*

That is, the affine hull of C is defined by $x + span(C - C)$ for any fixed $x \in C$. If C is convex, then it is straightforward to show that

$$aff(C) = \{\lambda x + (1 - \lambda)y : x, y \in C \text{ and } \lambda \in \mathbb{R}\}. \tag{8}$$

Intuitively, if we drew a line through any two points of C, then that entire line would necessarily be included in any affine subspace that contains C.

We will also replace the continuity assumption with the more restrictive assumption of Lipschitz continuity. For a convex subset C of X, a function $f : C \to \mathbb{R}$ is said to be *Lipschitz continuous* if there is some real number K such that for every $x, y \in C$, $|f(x) - f(y)| \leq K\|x - y\|$. The number K is called the *Lipschitz constant* of f.

We are now ready to state our variation of Mazur's theorem.

Proposition 1 *Suppose X is a separable Banach space and C is a closed and convex subset of X containing the origin, and suppose $aff(C)$ is dense in X. If $f : C \to \mathbb{R}$ is Lipschitz continuous and convex, then the set of points x where $\partial f(x)$ is a singleton is a dense G_δ (in the relative topology) set in C.*

Before proving this proposition, we will show that under the assumptions of the proposition, $\partial f(x) \neq \emptyset$ for all $x \in C$.

Lemma 1 *Suppose C is a convex subset of a Banach space X. If $f : C \to \mathbb{R}$ is Lipschitz continuous and convex, then $\partial f(x) \neq \emptyset$ for all $x \in C$. In particular, if $K \geq 0$ is the Lipschitz constant of f, then for all $x \in C$ there exists $x^* \in \partial f(x)$ with $\|x^*\| \leq K$.*

Proof of Lemma 1: We begin by introducing the standard definition of the epigraph of a function $f : C \to \mathbb{R}$:

$$epi(f) = \{(x, t) \in C \times \mathbb{R} : t \geq f(x)\}.$$

[5] See Phelps (1993, Theorem 1.20). An equivalent characterization in terms of closed convex sets and smooth points can be found in Holmes (1975, p. 171).

Note that $epi(f) \subset X \times \mathbb{R}$ is a convex set because f is convex with a convex domain C. Now, define

$$H = \{(x,t) \in X \times \mathbb{R} : t < -K\|x\|\}.$$

It is easily seen that H is nonempty and convex. Also, since $\|\cdot\|$ is necessarily continuous, H is open (in the product topology).

Let $x \in C$ be arbitrary. Let $H(x)$ be the translate of H so that its vertex is $(x, f(x))$; that is, $H(x) = (x, f(x)) + H$. We claim that $epi(f) \cap H(x) = \emptyset$. To see this, note first that

$$\begin{aligned} H(x) &= \{(x+y, f(x)+t) \in X \times \mathbb{R} : t < -K\|y\|\} \\ &= \{(y,t) \in X \times \mathbb{R} : t < f(x) - K\|y-x\|\}. \end{aligned}$$

Now, suppose $(y,t) \in epi(f)$, so that $t \geq f(y)$. By Lipschitz continuity, we have $f(y) \geq f(x) - K\|y-x\|$. Therefore, $t \geq f(x) - K\|y-x\|$, which implies $(y,t) \notin H(x)$.

Since $H(x)$ is open and nonempty, it has an interior point. We have also shown that $H(x)$ and $epi(f)$ are disjoint convex sets. Therefore, a version of the Separating Hyperplane Theorem implies there exists a nonzero continuous linear functional $(x^*, \lambda) \in X^* \times \mathbb{R}$ that separates $H(x)$ and $epi(f)$.[6] That is, there exists a scalar δ such that

$$\langle y, x^* \rangle + \lambda t \geq \delta \quad \text{if} \quad (y,t) \in epi(f) \tag{9}$$

and

$$\langle y, x^* \rangle + \lambda t \leq \delta \quad \text{if} \quad (y,t) \in H(x). \tag{10}$$

Clearly, we cannot have $\lambda < 0$. Also, if $\lambda = 0$, then Equation (10) implies $x^* = 0$. This would contradict (x^*, λ) being a nonzero functional. Therefore, $\lambda > 0$. Without loss of generality, we can take $\lambda = 1$, for otherwise we could renormalize (x^*, λ) by dividing by λ.

Since $(x, f(x)) \in epi(f)$, we have $\langle x, x^* \rangle + f(x) \geq \delta$. For all $t > 0$, we have $(x, f(x) - t) \in H(x)$, which implies $\langle x, x^* \rangle + f(x) - t \leq \delta$. Therefore, $\langle x, x^* \rangle + f(x) = \delta$, and thus for all $y \in C$,

$$\langle y, x^* \rangle + f(y) \geq \delta = \langle x, x^* \rangle + f(x).$$

Equivalently, we can write $f(y) - f(x) \geq \langle x - y, x^* \rangle = \langle y - x, -x^* \rangle$. Thus, $-x^* \in \partial f(x)$.

It remains only to show that $\|-x^*\| = \|x^*\| \leq K$. Suppose to the contrary. Then, there exists $y \in X$ such that $\langle y, x^* \rangle > K\|y\|$, and hence there also exists $\epsilon > 0$ such that $\langle y, x^* \rangle - \epsilon > K\|y\|$. Therefore,

$$\langle y + x, x^* \rangle + f(x) - K\|y\| - \epsilon > \langle x, x^* \rangle + f(x) = \delta,$$

which, by Equation (10), implies $(y + x, f(x) - K\|y\| - \epsilon) \notin H(x)$. However, this contradicts the definition of $H(x)$. Thus $\|-x^*\| \leq K$. ■

[6]See Aliprantis and Border (1999, Theorem 5.50) or Luenberger (1969, p. 133).

We are now ready to prove our main proposition.

Proof of Proposition 1: This proof is a variation of the proof of Mazur's theorem found in Phelps (1993). Since any subset of a separable Banach space is separable, *aff* (C) is separable. Let $\{x_n\} \subset$ *aff* (C) be a sequence which is dense in *aff* (C), and hence, by the density of *aff* (C) in X, also dense in X. Let K be the Lipschitz constant of f. For each $m, n \in \mathbb{N}$, let $A_{m,n}$ denote the set of $x \in C$ for which there exist $x^*, y^* \in \partial f(x)$ such that $\|x^*\|, \|y^*\| \leq 2K$ and

$$\langle x_n, x^* - y^* \rangle \geq \tfrac{1}{m}.$$

We claim that if $\partial f(x)$ is not a singleton for $x \in C$, then $x \in A_{m,n}$ for some $m, n \in \mathbb{N}$. By Lemma 1, for all $x \in C$, $\partial f(x) \neq \emptyset$. Therefore, if $\partial f(x)$ is not a singleton, then there exist $x^*, y^* \in \partial f(x)$ such that $x^* \neq y^*$. This does not tell us anything about the norm of x^* or y^*, but by Lemma 1, there exists $z^* \in \partial f(x)$ such that $\|z^*\| \leq K$. Either $z^* \neq x^*$ or $z^* \neq y^*$, so it is without loss of generality that we assume the former. It is straightforward to verify that the subdifferential is convex. Therefore, for all $\lambda \in (0, 1)$, $\lambda x^* + (1 - \lambda)z^* \in \partial f(x)$, and

$$\|\lambda x^* + (1 - \lambda)z^*\| \leq \|z^*\| + \lambda\|x^* - z^*\| \leq 2K$$

for λ sufficiently small. For some such λ, let $w^* = \lambda x^* + (1 - \lambda)z^*$. Then, $w^* \neq z^*$ and $\|w^*\| \leq 2K$. Since $w^* \neq z^*$, there exists $y \in X$ such that $\langle y, w^* - z^* \rangle > 0$. By the continuity of $w^* - z^*$, there exists a neighborhood N of y such that for all $z \in N$, $\langle z, w^* - z^* \rangle > 0$. Since $\{x_n\}$ is dense in X, there exists $n \in \mathbb{N}$ such that $x_n \in N$. Thus $\langle x_n, w^* - z^* \rangle > 0$, and hence there exists $m \in \mathbb{N}$ such that $\langle x_n, w^* - z^* \rangle > \frac{1}{m}$. Therefore, $x \in A_{m,n}$.

We have just shown that the set of $x \in C$ for which $\partial f(x)$ is a singleton is $\bigcap_{m,n}(C \setminus A_{m,n})$. It remains only show that for each $m, n \in \mathbb{N}$, $C \setminus A_{m,n}$ is open (in the relative topology) and dense in C. Then, we can appeal to the Baire category theorem.

We first show that each $A_{m,n}$ is relatively closed. If $A_{m,n} = \emptyset$, then $A_{m,n}$ is obviously closed, so suppose otherwise. Consider any sequence $\{z_k\} \subset A_{m,n}$ such that $z_k \to z$ for some $z \in C$. We will show that $z \in A_{m,n}$. For each k, choose $x_k^*, y_k^* \in \partial f(z_k)$ such that $\|x_k^*\|, \|y_k^*\| \leq 2K$ and $\langle x_n, x_k^* - y_k^* \rangle \geq \frac{1}{m}$. Since X is separable, the closed unit ball of X^* endowed with the weak* topology is metrizable and compact, which implies any sequence in this ball has a weak*-convergent subsequence.[7] Therefore, the closed ball of radius $2K$ around the origin of X^* has this same property. Thus, without loss of generality, we can assume there exist $x^*, y^* \in X^*$ with $\|x^*\|, \|y^*\| \leq 2K$ such that $x_k^* \xrightarrow{w^*} x^*$ and $y_k^* \xrightarrow{w^*} y^*$. Therefore, for any

[7] For metrizability, see Aliprantis and Border (1999, Theorem 6.34). Compactness follows from Alaoglu's theorem; see Aliprantis and Border (1999, Theorem 6.25). Note that compactness only guarantees that every net has a convergent subnet, but compactness and metrizability together imply that every sequence has a convergent subsequence.

$y \in C$, we have

$$\langle y - z, x^* \rangle = \lim_{k\to\infty} \langle y - z_k, x_k^* \rangle \leq \lim_{k\to\infty} \big[f(y) - f(z_k)\big] = f(y) - f(z),$$

which implies $x^* \in \partial f(z)$. A similar argument shows $y^* \in \partial f(z)$. Finally, since

$$\langle x_n, x^* - y^* \rangle = \lim_{k\to\infty} \langle x_n, x_k^* - y_k^* \rangle \geq \tfrac{1}{m},$$

we have $z \in A_{m,n}$, and hence $A_{m,n}$ is relatively closed.

We now need to show that $C \setminus A_{m,n}$ is dense in C for each $m, n \in \mathbb{N}$. Consider arbitrary $m, n \in \mathbb{N}$ and $z \in C$. We will find a sequence $\{z_k\} \subset C \setminus A_{m,n}$ such that $z_k \to z$. Since C contains the origin, $\mathit{aff}(C)$ is a subspace of X. Hence, $z + x_n \in \mathit{aff}(C)$, so Equation (8) implies there exist $x, y \in C$ and $\lambda \in \mathbb{R}$ such that $\lambda x + (1-\lambda)y = z + x_n$. Let us first suppose $\lambda > 1$; we will consider the other cases shortly. Note that $\lambda > 1$ implies $0 < \frac{\lambda-1}{\lambda} < 1$. Consider any sequence $\{a_k\} \subset (0, \frac{\lambda-1}{\lambda})$ such that $a_k \to 0$. Define a sequence $\{y_k\} \subset C$ by $y_k = a_k y + (1-a_k)z$, and note that $y_k \to z$. We claim that for each $k \in \mathbb{N}$, $y_k + \frac{a_k}{\lambda-1}x_n \in C$. To see this, note the following:

$$\begin{aligned}
y_k + \tfrac{a_k}{\lambda-1}x_n &= a_k y + (1-a_k)z + \tfrac{a_k}{\lambda-1}(x_n + z - z) \\
&= a_k y + (1-a_k)z + \tfrac{a_k}{\lambda-1}\big(\lambda x + (1-\lambda)y - z\big) \\
&= (1-a_k)z + \tfrac{a_k\lambda}{\lambda-1}x - \tfrac{a_k}{\lambda-1}z \\
&= \big(1 - \tfrac{a_k\lambda}{\lambda-1}\big)z + \tfrac{a_k\lambda}{\lambda-1}x
\end{aligned}$$

Since $0 < a_k < \frac{\lambda-1}{\lambda}$, we have $0 < \frac{a_k\lambda}{\lambda-1} < 1$. Thus $y_k + \frac{a_k}{\lambda-1}x_n$ is a convex combination of z and x, so it is an element of C.

Consider any $k \in \mathbb{N}$. Because C is convex, we have $y_k + t x_n \in C$ for all $t \in (0, \frac{a_k\lambda}{\lambda-1})$. Define a function $g : (0, \frac{a_k\lambda}{\lambda-1}) \to \mathbb{R}$ by $g(t) = f(y_k + t x_n)$, and note that g is convex. It is a standard result that a convex function on an open interval in $\mathbb{R}$ is differentiable for all but (at most) countably many points of this interval.[8] Let t_k be any $t \in (0, \frac{a_k\lambda}{\lambda-1})$ at which $g'(t)$ exists, and let $z_k = y_k + t_k x_n$. If $x^* \in \partial f(z_k)$, then it is straightforward to verify that the linear mapping $t \mapsto t\langle x_n, x^* \rangle$ is a subdifferential to g at t_k. Since g is differentiable at t_k, it can only have one element in its subdifferential at that point. Therefore, for any $x^*, y^* \in \partial f(z_k)$, we have $\langle x_n, x^* \rangle = \langle x_n, y^* \rangle$, and hence $z_k \in C \setminus A_{m,n}$. Finally, note that since $0 < t_k < \frac{a_k\lambda}{\lambda-1}$ and $a_k \to 0$, we have $t_k \to 0$. Therefore, $z_k = y_k + t_k x_n \to z$.

We did restrict attention above the case of $\lambda > 1$. However, if $\lambda < 0$, then let $\lambda' = 1 - \lambda > 1$, $x' = y$, $y' = x$, and the analysis is the same as above. If $\lambda \in [0,1]$, then note that $z + x_n \in C$. Similar to in the preceding paragraph, for any $k \in \mathbb{N}$, define a function $g : (0, \frac{1}{k}) \to \mathbb{R}$ by $g(t) = f(z + t x_n)$. Let t_k be any $t \in (0, \frac{1}{k})$ at which $g'(t)$ exists, and let $z_k = z + t_k x_n$. Then, as argued above, $z_k \in C \setminus A_{m,n}$ for all $k \in \mathbb{N}$ and $z_k \to z$.

[8] See Phelps (1993, Theorem 1.16).

We have now proved that for each $m, n \in \mathbb{N}$, $C \setminus A_{m,n}$ is open (in the relative topology) and dense in C. Since C is a closed subset of a Banach space, it is a Baire space, which implies every countable intersection of (relatively) open dense subsets of C is also dense.[9] This completes the proof. ■

B Proof of Theorem 1

The necessity of the axioms in Theorem 1 is straightforward and left to the reader. For the sufficiency direction, let $\mathcal{A}^c \subset \mathcal{A}$ denote the set of convex menus. In both parts A and B of Theorem 1, $\succsim$ satisfies IR. In part A, IR is directly assumed whereas in part B it is implied by weak order, continuity, MON, and ACP (see Lemma 2). Therefore for all $A \in \mathcal{A}$, $A \sim co(A) \in \mathcal{A}^c$. Note that for any $u \in \mathcal{U}$, we have

$$\max_{p \in A} u \cdot p = \max_{p \in co(A)} u \cdot p.$$

Thus it is enough to establish the representations in Theorem 1 for convex menus and then define the utility of $A \notin \mathcal{A}^c$ to be that of $co(A)$. The representations in Theorem 1 for convex menus are implied by the results in Propositions 2, 3, and 4 which are presented respectively in sections B.2, B.3, and B.4.

B.1 Preliminary Observations

In this section we establish a number of simple implications of the axioms introduced in the text. These results will be useful in subsequent sections.

Lemma 2 *If $\succsim$ satisfies weak order, ACP, MON, and continuity, then it also satisfies IR.*

Proof: Let $A \in \mathcal{A}$. Monotonicity implies that $co(A) \succsim A$, hence we only need to prove that $A \succsim co(A)$. Let us inductively define a sequence of sets via $A_0 = A$ and $A_k = \frac{1}{2}A_{k-1} + \frac{1}{2}A_{k-1}$ for $k \geq 1$. ACP implies that $A_{k-1} \succsim A_k$, therefore by transitivity $A \succsim A_k$ for any k. It is straightforward to verify that $d_h(A_k, co(A)) \to 0$, so we have $A \succsim co(A)$ by continuity. ■

We do not assume that independence holds on $\mathcal{A}$, but we will see that our other axioms imply that independence does hold for singletons.

Axiom 7 (Singleton Independence): *For all $p, q, r \in \triangle(Z)$ and $0 < \lambda < 1$,*

$$\{p\} \succsim \{q\} \iff \lambda\{p\} + (1-\lambda)\{r\} \succsim \lambda\{q\} + (1-\lambda)\{r\}.$$

[9] See Theorems 3.34 and 3.35 of Aliprantis and Border (1999).

Lemma 3 *If $\succsim$ satisfies weak order, continuity, and TI, then $\succsim$ must also satisfy singleton independence.*

Proof: In fact, if $\succsim$ satisfies weak order, continuity, and either TI-1 or TI-2, then it must satisfy singleton independence. We will prove the result using TI-1 and leave the alternative proof using TI-2 to the reader.

First, we show that for any $\lambda \in (0,1)$,

$$\{p\} \succsim \{q\} \iff \{p\} \succsim (1-\lambda)\{p\} + \lambda\{q\}. \tag{11}$$

This can be proved using a simple induction argument. Note that if

$$\left(1 - \tfrac{m-1}{n}\right)\{p\} + \left(\tfrac{m-1}{n}\right)\{q\} \succsim \left(1 - \tfrac{m}{n}\right)\{p\} + \left(\tfrac{m}{n}\right)\{q\},$$

for $m, n \in \mathbb{N}$, $m < n$, then adding $\theta = \frac{1}{n}(q-p)$ to each side and applying TI-1 implies

$$\left(1 - \tfrac{m}{n}\right)\{p\} + \left(\tfrac{m}{n}\right)\{q\} \succsim \left(1 - \tfrac{m+1}{n}\right)\{p\} + \left(\tfrac{m+1}{n}\right)\{q\},$$

Now suppose that $\{p\} \succsim \left(1 - \frac{1}{n}\right)\{p\} + \left(\frac{1}{n}\right)\{q\}$. Then, using induction and the transitivity of $\succsim$, we obtain the following:

$$\{p\} \succsim \left(1 - \tfrac{1}{n}\right)\{p\} + \left(\tfrac{1}{n}\right)\{q\} \succsim \cdots \succsim \left(\tfrac{1}{n}\right)\{p\} + \left(1 - \tfrac{1}{n}\right)\{q\} \succsim \{q\}. \tag{12}$$

A similar line of reasoning shows that if $\{p\} \prec \left(1 - \frac{1}{n}\right)\{p\} + \left(\frac{1}{n}\right)\{q\}$, then we obtain the following:[10]

$$\{p\} \prec \left(1 - \tfrac{1}{n}\right)\{p\} + \left(\tfrac{1}{n}\right)\{q\} \prec \cdots \prec \left(\tfrac{1}{n}\right)\{p\} + \left(1 - \tfrac{1}{n}\right)\{q\} \prec \{q\}. \tag{13}$$

In sum, Equations (12) and (13) imply that for any $m, n \in \mathbb{N}$, $m < n$, we have

$$\{p\} \succsim \{q\} \iff \{p\} \succsim \left(1 - \tfrac{1}{n}\right)\{p\} + \left(\tfrac{1}{n}\right)\{q\} \iff \{p\} \succsim \left(1 - \tfrac{m}{n}\right)\{p\} + \left(\tfrac{m}{n}\right)\{q\}$$

This establishes the claim in Equation (11) for $\lambda \in (0,1) \cap \mathbb{Q}$. The continuity of $\succsim$ implies that Equation (11) holds for all $\lambda \in (0,1)$.

[10] Note that here we need the converse of TI-1: $A \succ B \Rightarrow A + \theta \succ B + \theta$. However, this relationship is implied by TI-1, for suppose $B + \theta \succsim A + \theta$. Then, by TI-1,

$$B = (B + \theta) + (-\theta) \succsim (A + \theta) + (-\theta) = A.$$

Finally, for any $\lambda \in (0,1)$, take $\theta = (1-\lambda)(r-p)$. Then, $\{p\}+\theta = \lambda\{p\}+(1-\lambda)\{r\}$ and $(1-\lambda)\{p\}+\lambda\{q\}+\theta = \lambda\{q\}+(1-\lambda)\{r\}$. Therefore, by Equation (11) and TI-1,

$$\begin{aligned}
\{p\} \succsim \{q\} &\iff \{p\} \succsim (1-\lambda)\{p\}+\lambda\{q\} \\
&\iff \lambda\{p\}+(1-\lambda)\{r\} \succsim \lambda\{q\}+(1-\lambda)\{r\},
\end{aligned}$$

so singleton independence is satisfied. ■

The next axiom is a basic technical requirement.

Axiom 8 (Singleton Nontriviality): *There exist $p, q \in \triangle(Z)$ such that $\{p\} \succ \{q\}$.*

We explained in our discussion of the representation that strong continuity implies singleton nontriviality. In the following sections, we will also establish a counterpart of our representation result by replacing strong continuity with the weaker assumptions of continuity and singleton nontriviality. We conclude this section by stating some other useful lemmas relating to translation invariance.

Lemma 4 *If $\succsim$ satisfies weak order, continuity, and TI-2, then $\succsim$ must also satisfy the following condition: Suppose $A \in \mathcal{A}$, $\theta \in \Theta$, $A+\theta \in \mathcal{A}$, and $A+\theta \succ A$. Then, for all $A' \in \mathcal{A}$ and $k>0$,*

$$\begin{aligned}
A'+k\theta \in \mathcal{A} &\Rightarrow A'+k\theta \succ A' \\
A'-k\theta \in \mathcal{A} &\Rightarrow A'-k\theta \prec A'
\end{aligned}$$

Proof: By TI-2, for any $q \in \triangle(Z)$,

$$\{q\}+\theta \in \mathcal{A} \quad \Rightarrow \quad \{q\}+\theta \succ \{q\}.$$

Combining this result with singleton independence (see Lemma 3) implies that for all $q \in \triangle(Z)$ and $k>0$,

$$\begin{aligned}
\{q\}+k\theta \in \mathcal{A} &\Rightarrow \{q\}+k\theta \succ \{q\} \\
\{q\}-k\theta \in \mathcal{A} &\Rightarrow \{q\}-k\theta \prec \{q\}
\end{aligned}$$

Applying TI-2 again implies that for all $A' \in \mathcal{A}$ and $k>0$,

$$\begin{aligned}
A'+k\theta \in \mathcal{A} &\Rightarrow A'+k\theta \succ A' \\
A'-k\theta \in \mathcal{A} &\Rightarrow A'-k\theta \prec A'
\end{aligned}$$

■

Lemma 5 *If $\succsim$ satisfies weak order, continuity, and TI, then $\succsim$ must also satisfy the following condition: Suppose $A, A' \in \mathcal{A}$ and $A \sim A'$, and suppose $\theta \in \Theta$ is such that there exists $B \in \mathcal{A}$ with $B + \theta \in \mathcal{A}$ and $B + \theta \succ B$. Then, for all $k, k' \in \mathbb{R}$ such that $A + k\theta, A + k'\theta \in \mathcal{A}$,*

$$A + k\theta \succ A' + k'\theta \quad \Longleftrightarrow \quad k > k'.$$

Proof: First, suppose $k > k'$. Then, $A + k'\theta \in \mathcal{A}$, and since $A \sim A'$, TI-1 implies $A + k'\theta \sim A' + k'\theta$. Since $k - k' > 0$, Lemma 4 implies

$$A + k\theta = A + k'\theta + (k - k')\theta \succ A + k'\theta \sim A' + k'\theta.$$

For the converse, suppose $k \le k'$. Then, $A' + k\theta \in \mathcal{A}$, and since $A \sim A'$, TI-1 implies $A + k\theta \sim A' + k\theta$. Since $k' - k \ge 0$, Lemma 4 implies

$$A' + k'\theta = A' + k\theta + (k' - k)\theta \succsim A' + k\theta \sim A + k\theta.$$

■

B.2 Construction of V

Note that $\mathcal{A}$ is a compact metric space since $\triangle(Z)$ is a compact metric space (see for instance Munkres, p279). It is a standard exercise to show that $\mathcal{A}^c$ is a closed subset of $\mathcal{A}$, hence $\mathcal{A}^c$ is also compact. Since linear functions on $\triangle(Z)$ are in a natural one-to-one correspondence with points in $\mathbb{R}^Z$, we will often use the notation $u \cdot p$ instead of $u(p)$ if u is a linear function on $\triangle(Z)$ and $p \in \triangle(Z)$. Remember that for any metric space (X, d), $f : X \to \mathbb{R}$ is *Lipschitz continuous* if there is some real number K such that for every $x, y \in X$, $|f(x) - f(y)| \le K d(x, y)$. The number K is called the *Lipschitz constant* of f.

We will construct a function $V : \mathcal{A}^c \to \mathbb{R} \cup \{+\infty\}$ that represents $\succsim$ on $\mathcal{A}^c$ and has certain desirable properties. Let $\bar{\mathbb{R}}$ indicate the extended real numbers: $\bar{\mathbb{R}} \equiv \mathbb{R} \cup \{-\infty, +\infty\}$. We next define the notion of Θ-linearity in order to present the main result of this section.

Definition 6 *Suppose that $V : \mathcal{A}^c \to \bar{\mathbb{R}}$. Then V is* Θ-linear *if there exists $v \in \mathbb{R}^Z$ such that for all $A \in \mathcal{A}^c$ and $\theta \in \Theta$ with $A + \theta \in \mathcal{A}^c$, we have:*

$$V(A + \theta) = V(A) + v \cdot \theta.$$

Proposition 2 *If the preference $\succsim$ satisfies weak order, continuity, ACP, TI, and singleton nontriviality, then there exists a function $V : \mathcal{A}^c \to \mathbb{R} \cup \{+\infty\}$ with the following properties:*

1. *V is continuous, convex, and Θ-linear.*
2. *For any $A, B \in \mathcal{A}^c$: $A \succsim B \iff V(A) \ge V(B)$.*

3. There exist $p, q \in \triangle(Z)$ such that $V(\{p\}) > V(\{q\})$.

4. The function $v : \triangle(Z) \to \mathbb{R}$ defined by $v(p) = V(\{p\})$ is in $\mathcal{U}$.

If $\succsim$ also satisfies strong-continuity then V above is Lipschitz continuous.

Note that part 3 will follow immediately from singleton nontriviality once we show part 2. In the remainder of this section assume that $\succsim$ satisfies all the axioms in Proposition 2 except strong continuity. We will explicitly assume strong continuity of $\succsim$ at the end of the section in Lemma 13, in order to prove Lipschitz continuity of V.

Let $\mathcal{S} \equiv \{\{q\} : q \in \triangle(Z)\}$ be the set all of singleton sets in $\mathcal{A}^c$. Given the assumptions of Proposition 2 and the results of Lemma 3, $\succsim$ satisfies the von Neumann-Morgenstern axioms on $\mathcal{S}$. Therefore, there exists $v \in \mathbb{R}^Z$ such that for all $p, q \in \triangle(Z)$, $\{p\} \succsim \{q\}$ iff $v \cdot p \geq v \cdot q$. Since $\succsim$ satisfies singleton nontriviality, we can assume without loss of generality that $v \in \mathcal{U}$. We will abuse notation and also treat v as a function $v : \mathcal{S} \to \mathbb{R}$ naturally defined by $v(\{p\}) = v \cdot p$.

Note that v is Θ-linear since $v(\{p\} + \theta) = v(\{p\}) + v \cdot \theta$ whenever $p \in \triangle(Z)$, $\theta \in \Theta$, and $p + \theta \in \triangle(Z)$. We want to extend v to a function on $\mathcal{A}^c$ that represents $\succsim$ and is Θ-linear. If it were the case that for all $A \in \mathcal{A}^c$ there exist $p, q \in \triangle(Z)$ such that $\{p\} \succsim A \succsim \{q\}$, then we could apply continuity to extend v in the desired way. However, this is not generally the case, so we must use a different method to extend v.

The outline of the construction of the desired V is the following. We have already defined a function, v, that represents $\succsim$ on $\mathcal{S}$ and is Θ-linear. We will construct a sequence of subsets of $\mathcal{A}^c$, starting with $\mathcal{S}$, such that each set is contained in its successor set. We will then extend v sequentially to each of these domains, while still representing $\succsim$ and preserving certain linearity properties. The domain will grow to eventually contain "almost all" of the sets in $\mathcal{A}^c$, and we show how to extend to all of $\mathcal{A}^c$. Finally, we prove that the function we construct is continuous, Θ-linear, and convex.

Before proceeding, let $p^*, p_* \in \triangle(Z)$ be the most preferred and the least preferred elements in $\mathcal{S}$. That is, for all $p \in \triangle(Z)$, $\{p^*\} \succsim \{p\} \succsim \{p_*\}$. Singleton Nontriviality implies that $\{p^*\} \succ \{p_*\}$. Define $\theta^* = p^* - p_*$.

Consider a sequence of subsets (of $\mathcal{A}^c$), $\mathcal{A}_0, \mathcal{A}'_0, \mathcal{A}_1, \mathcal{A}'_1, \ldots$, defined as follows: Let $\mathcal{A}_0 \equiv \mathcal{S}$. Define $\mathcal{A}_i$ for all $i \geq 1$ by

$$\mathcal{A}_i \equiv \{A \in \mathcal{A}^c : A = B + k\theta^* \text{ for some } k \in \mathbb{R}, B \in \mathcal{A}'_{i-1}\}.$$

Define $\mathcal{A}'_i$ for all $i \geq 0$ by

$$\mathcal{A}'_i \equiv \{A \in \mathcal{A}^c : A \sim B \text{ for some } B \in \mathcal{A}_i\}.$$

Intuitively, we first extend $\mathcal{S}$ by including all $A \in \mathcal{A}^c$ that are viewed with indifference to some $B \in \mathcal{S}$. Then we extend to all translations by multiples of θ^*. We repeat the process, alternating between extension by indifference and extension by translation. Note that $\mathcal{S} = \mathcal{A}_0 \subset \mathcal{A}'_0 \subset \mathcal{A}_1 \subset \mathcal{A}'_1 \subset \cdots$.

We also define a sequence of functions, $V_0, V_0', V_1, V_1', \ldots$, from these domains. That is, for all $i \geq 0$, $V_i : \mathcal{A}_i \to \mathbb{R}$ and $V_i' : \mathcal{A}_i' \to \mathbb{R}$. Define these functions recursively as follows: First, let $V_0 = v$. Then, for $i \geq 0$, if $A \in \mathcal{A}_i'$, then $A \sim B$ for some $B \in \mathcal{A}_i$, so define V_i' by $V_i'(A) = V_i(B)$. For $i \geq 1$, if $A \in \mathcal{A}_i$, then $A = B + k\theta^*$ for some $k \in \mathbb{R}$ and $B \in \mathcal{A}_{i-1}'$, so define V_i by $V_i(A) = V_{i-1}'(B) + v \cdot k\theta^*$. In a series of lemmas, we will show that these are well-defined functions which represent $\succsim$ on their domains and are θ^*-linear. By θ^*-linearity, we mean that $V(A + k\theta^*) = V(A) + v \cdot k\theta^*$, but this equality may not hold for other $\theta \in \Theta$ (although we will later prove that it does).

First, we present a useful result regarding $\mathcal{A}_i'$.

Lemma 6 *For any $i \geq 0$, if $A, B \in \mathcal{A}_i'$ and $A \precsim C \precsim B$, then $C \in \mathcal{A}_i'$.*

Proof: We proceed by induction on i. To prove the result for $\mathcal{A}_0'$, suppose $A, B \in \mathcal{A}_0'$ and $A \precsim C \precsim B$ for some $C \in \mathcal{A}^c$. Since $A, B \in \mathcal{A}_0'$, there exist $p, q \in \triangle(Z)$ such that $\{p\} \sim A \precsim C \precsim B \sim \{q\}$. Continuity implies there exists a $\lambda \in [0, 1]$ such that $\{\lambda p + (1 - \lambda)q\} \sim C$. By definition of $\mathcal{A}_0'$, this implies that $C \in \mathcal{A}_0'$.

We now show that if $\mathcal{A}_{i-1}'$ satisfies the desired condition for $i \geq 1$, then $\mathcal{A}_i'$ does also. Suppose $A, B \in \mathcal{A}_i'$ and $A \precsim C \precsim B$ for some $C \in \mathcal{A}^c$. If there exist $A', B' \in \mathcal{A}_{i-1}'$ such that $A' \precsim C \precsim B'$, then $C \in \mathcal{A}_{i-1}' \subset \mathcal{A}_i'$ by the induction assumption. Thus, WLOG, suppose $C \succ A'$ for all $A' \in \mathcal{A}_{i-1}'$. Since $B \in \mathcal{A}_i'$, there exists a $B' \in \mathcal{A}_i$ such that $B' \sim B \succsim C$. Since $B' \in \mathcal{A}_i$, there exists a $A' \in \mathcal{A}_{i-1}'$ and $k \in \mathbb{R}$ such that $B' = A' + k\theta^*$. Since $A' \in \mathcal{A}_{i-1}'$ implies $C \succ A'$, this implies $A' + k\theta^* \succsim C \succ A'$. By continuity, there exists a $k' \in [0, k]$ such that $A' + k'\theta^* \sim C$. But $A' + k'\theta^* \in \mathcal{A}_i$, so it must be that $C \in \mathcal{A}_i'$. Therefore, by induction, the desired condition is satisfied for all $i \geq 0$. ■

We now prove the desired properties of V_i and V_i'.

Lemma 7 *For all $i \geq 0$, V_i and V_i' are well-defined functions which represent $\succsim$ on their domains and are θ^*-linear.*

Proof: We again proceed by induction on i. Obviously, V_0 satisfies the desired conditions. We show that for all $i \geq 0$, if V_i satisfies the desired conditions, then so must V_i'. Then, we show that for all $i \geq 1$, if V_{i-1}' satisfies the desired conditions, then so must V_i.

Suppose V_i satisfies the desired conditions for some $i \geq 0$. We need to show that V_i' does also. The transitivity of $\succsim$ implies that V_i' is well-defined and represents $\succsim$ on $\mathcal{A}_i'$. For suppose $A \in \mathcal{A}_i'$ and $B, B' \in \mathcal{A}_i$ are such that $A \sim B$ and $A \sim B'$. By the induction assumption, $V_i(B) = V_i(B')$, so $V_i'(A)$ is uniquely defined. Also, if $A, A' \in \mathcal{A}_i'$, then there exist $B, B' \in \mathcal{A}_i$ such that $B \sim A$ and $A' \sim B'$. Therefore, $V_i'(A) = V_i(B) \geq V_i(B') = V_i'(A')$ iff $B \succsim B'$ iff $A \succsim A'$, so V_i' represents $\succsim$.

We now need to show that V_i' is θ^*-linear. First, we will prove this for V_0'. Suppose $A, A+k\theta^* \in \mathcal{A}_0'$ where without loss of generality $k > 0$.[11] By Lemma 4, this implies $A+k\theta^* \succ A$. Since $A + k\theta^* \in \mathcal{A}_0'$, there exists a $q \in \triangle(Z)$ such that $A + k\theta^* \sim \{q\}$. Now, since $\{p_*\}+\theta^* = \{p^*\} \succsim \{q\} \succsim \{p_*\}$, there exists a $k' \in [0,1]$ such that $\{p_*\}+k'\theta^* \sim \{q\} \sim A+k\theta^*$. Similarly, there exists a $k'' \in [0,1]$ such that $\{p_*\} + k''\theta^* \sim A$. Therefore, by Lemma 5, $k' - k'' = k$, so that the θ^*-linearity of V_0 implies

$$\begin{aligned} V_0'(A+k\theta^*) - V_0'(A) &= V_0(\{p_*\}+k'\theta^*) - V_0(\{p_*\}+k''\theta^*) \\ &= V_0(\{p_*\}) + v\cdot k'\theta^* - V_0(\{p_*\}) - v \cdot k''\theta^* \\ &= v\cdot(k'-k'')\theta^* \ = \ v\cdot k\theta^*. \end{aligned}$$

Thus V_0' is θ^*-linear.

We now prove that for all $i \geq 1$, if V_i satisfies the desired conditions, then V_i' is θ^*-linear. Suppose $A, A+k\theta^* \in \mathcal{A}_i'$ where without loss of generality $k > 0$. First, consider the case of $A+k\theta^* \succsim C \succsim A$ for some $C \in \mathcal{A}_{i-1}'$. By continuity, this implies there exists a $k' \in [0,k]$ such that $A + k'\theta^* \sim C$. However, this implies $A + k'\theta^* \in \mathcal{A}_{i-1}'$, which implies $A, A+k\theta^* \in \mathcal{A}_i$. Since V_i' is equal to V_i on $\mathcal{A}_i$ and V_i is assumed to satisfy θ^*-linearity, we have the desired result, $V_i'(A+k\theta^*) = V_i'(A) + v\cdot k\theta^*$. Now, we will consider the case of $A+k\theta^* \succ A \succ C$ for all $C \in \mathcal{A}_{i-1}'$. (The proof for $C \succ A+k\theta^* \succ A$ for all $C \in \mathcal{A}_{i-1}'$ is similar.) Since $A+k\theta^* \in \mathcal{A}_i'$, there exists a $B \in \mathcal{A}_i$ such that $A + k\theta^* \sim B$. If $B - k\theta^* \in \mathcal{A}^c$, the we can use TI-1 and the θ^*-linearity of V_i to obtain the desired result. However, it is not obvious that $B - k\theta^* \in \mathcal{A}^c$. Since $B \in \mathcal{A}_i$, there exists a $C \in \mathcal{A}_{i-1}'$ and a $k' \in \mathbb{R}$ such that $B = C + k'\theta^*$. Thus, because $C \in \mathcal{A}_{i-1}'$, we have $C + k'\theta^* \succ A \succ C$, so by continuity there exists a $k'' \in [0,k']$ such that $C + k''\theta^* \sim A$. Now, $C + k''\theta^* \sim A$ and $C + k'\theta^* \sim A + k\theta^*$, so by Lemma 5, $k' - k'' = k$. Therefore,

$$\begin{aligned} V_i'(A+k\theta^*) - V_i'(A) &= V_i(C+k'\theta^*) - V_i(C+k''\theta^*) \\ &= V_i(C) + v\cdot k'\theta^* - V_i(C) - v\cdot k''\theta^* \\ &= v\cdot(k'-k'')\theta^* \ = \ v\cdot k\theta^*, \end{aligned}$$

so θ^*-linearity is satisfied.

We now show that for all $i \geq 1$, if V_{i-1}' satisfies the desired conditions, then so must V_i. First, note that the extension from $V_{i-1}' : \mathcal{A}_{i-1}' \to \mathbb{R}$ to $V_i : \mathcal{A}_i \to \mathbb{R}$ is well-defined. For suppose $A \in \mathcal{A}_i$ and $A = B + k\theta^* = B' + k'\theta^*$ for $B, B' \in \mathcal{A}_{i-1}'$. Then, $B = B' + (k'-k)\theta^*$. Since $B, B' \in \mathcal{A}_{i-1}'$ and V_{i-1}' is θ^*-linear, we have

$$V_{i-1}'(B) = V_{i-1}'(B') + v\cdot(k'-k)\theta^* = V_{i-1}'(B') + v\cdot k'\theta^* - v\cdot k\theta^*,$$

[11] Showing this for the case when $k > 0$ implies it for the case when $k < 0$: if $k < 0$ then $V(A) = V(A + k\theta^* - k\theta^*) = V(A+k\theta^*) - v\cdot k\theta^*$ so we again have $V(A+k\theta^*) = V(A) + v\cdot k\theta^*$.

which implies

$$V'_{i-1}(B) + v \cdot k\theta^* = V'_{i-1}(B') + v \cdot k'\theta^*.$$

Thus $V_i(A)$ is uniquely defined.

To see that V_i is θ^*-linear, suppose $A \in \mathcal{A}_i$ and $A' = A + k\theta^* \in \mathcal{A}_i$. Then, $A = B + k'\theta^*$ for some $B \in \mathcal{A}'_{i-1}$, so $A' = B + (k' + k)\theta^*$. Therefore,

$$\begin{aligned} V_i(A') &= V'_{i-1}(B) + v \cdot (k + k')\theta^* \\ &= V'_{i-1}(B) + v \cdot k\theta^* + v \cdot k'\theta^* \\ &= V_i(A) + v \cdot k'\theta^*. \end{aligned}$$

To see that V_i represents $\succsim$ on $\mathcal{A}_i$, suppose $A, A' \in \mathcal{A}_i$. Therefore, $A = B + k\theta^*$ and $A' = B' + k'\theta^*$ for some $B, B' \in \mathcal{A}'_{i-1}$, $k, k' \in \mathbb{R}$. If $k = k'$, then by TI-1 and the definition of V_i, $A \succsim A'$ iff $B \succsim B'$ iff $V'_{i-1}(B) \geq V'_{i-1}(B')$ iff $V_i(A) \geq V_i(A')$. However, it may not be the case that $k = k'$. There are several possibilities when $k \neq k'$. We work through one of them here: $A \succsim B' \succsim B$ and $A' \succsim B' \succsim B$. The other cases are similar. Notice that we have $B + k\theta^* \succsim B' \succsim B$. This implies that $k \geq 0$ (see Lemma 4), and continuity implies there exists a $k'' \in [0, k]$ such that $B + k''\theta^* \sim B'$. Let $C = B + k''\theta^* \in \mathcal{A}'_{i-1}$. Then, we have $C \sim B'$, $A = C + (k - k'')\theta^*$, and $A' = B' + k'\theta^*$. Given Lemma 5, this requires that $A \succsim A'$ iff $k - k'' \geq k'$ iff

$$\begin{aligned} V_i(A) &= V'_{i-1}(C) + v \cdot (k - k'')\theta^* \\ &= V'_{i-1}(B') + v \cdot (k - k'')\theta^* \\ &\geq V'_{i-1}(B') + v \cdot k'\theta^* \;=\; V_i(A'). \end{aligned}$$

By induction, we see that for all $i \geq 0$, V_i and V'_i are well-defined, represent $\succsim$ on their respective domains, and satisfy θ^*-linearity. ■

We can define a function $\hat{V} : \bigcup_i \mathcal{A}_i \to \mathbb{R}$ by $\hat{V}(A) = V_i(A)$ if $A \in \mathcal{A}_i$. This is well-defined because if $A \in \mathcal{A}_i$ and $A \in \mathcal{A}_{i'}$, then WLOG suppose $\mathcal{A}_i \subset \mathcal{A}_{i'}$. Then $V_{i'}(B) = V_i(B)$ for all $B \in \mathcal{A}_i$, so that $V_{i'}(A) = V_i(A)$. Note that $\hat{V}$ represents $\succsim$ on $\bigcup_i \mathcal{A}_i$ and is θ^*-linear.

It is possible that $\bigcup_i \mathcal{A}_i = \mathcal{A}^c$, but this is not necessarily the case. We will now define a subset of $\bigcup_i \mathcal{A}_i$ that will be very useful in extending $\hat{V}$ to all of $\mathcal{A}^c$ and also in proving certain properties of the function we construct. Define the set $\mathcal{I} \subset \mathcal{A}^c$ as follows.

$$\mathcal{I} \equiv \{A \in \mathcal{A}^c : \forall \theta \in \Theta \; \exists k > 0 \text{ such that } A + k\theta \in \mathcal{A}^c\}. \tag{14}$$

Thus $\mathcal{I}$ contains menus that can be translated at least a "little bit" in the direction of any vector in Θ. The following lemma will be helpful in determining exactly what part of $\mathcal{A}^c$ is contained in $\bigcup_i \mathcal{A}_i$.

Lemma 8 $\mathcal{I} \subset \bigcup_i \mathcal{A}_i$.

Proof: Consider any set $A \in \mathcal{I}$. By the definition of $\mathcal{I}$, there exists some $k > 0$ such that $A + k\theta^*, A - k\theta^* \in \mathcal{A}^c$. Now, choose any $q \in A$. Cleary, $\{q\} + k\theta^*, \{q\} - k\theta^* \in \mathcal{A}^c$. Now consider $B(\lambda) \equiv \lambda A + (1-\lambda)\{q\}$ for $\lambda \in [0,1]$. Note that $B(\lambda) + k\theta^*, B(\lambda) - k\theta^* \in \mathcal{A}^c$. This holds because

$$B(\lambda) + k\theta^* = \lambda A + (1-\lambda)\{q\} + k\theta^* = \lambda(A + k\theta^*) + (1-\lambda)(\{q\} + k\theta^*)$$

and similarly for $B(\lambda) - k\theta^*$. By Lemma 4, for all $\lambda \in [0,1]$, $B(\lambda) - k\theta^* \prec B(\lambda) \prec B(\lambda) + k\theta^*$. By continuity, for each λ there exists an open (relative to $[0,1]$) interval $e(\lambda)$ such that $\lambda \in e(\lambda)$ and for all λ' such that $\lambda' \in e(\lambda)$,

$$B(\lambda) - k\theta^* \prec B(\lambda') \prec B(\lambda) + k\theta^*.$$

Thus $\{e(\lambda) : \lambda \in [0,1]\}$ is an open cover of $[0,1]$. Since $[0,1]$ is compact, there exists a finite subcover, $\{e(\lambda_1), \ldots, e(\lambda_n)\}$. Assume the λ_i's are arranged so that $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$. That is, as i increases, $e(\lambda_i)$ moves "farther" from $\{q\}$ and "closer" to A. We can prove that $B(\lambda_1) \in \mathcal{A}_1$ by first observing that

$$B(\lambda_1) - k\theta^* \prec B(0) = \{q\} \prec B(\lambda_1) + k\theta^*,$$

which implies there exists $k' \in (-k, k)$ such that $B(\lambda_1) + k'\theta^* \sim \{q\}$. This implies $B(\lambda_1) + k'\theta^* \in \mathcal{A}'_0$, which implies that $B(\lambda_1) \in \mathcal{A}_1$. Now, we can show that $B(\lambda_i) \in \mathcal{A}_i$ implies $B(\lambda_{i+1}) \in \mathcal{A}_{i+1}$. If $B(\lambda_i) \in \mathcal{A}_i$, then we also have $B(\lambda_i) + k'\theta^* \in \mathcal{A}_i$ for all $k' \in (-k, k)$. Since $e(\lambda_i) \cap e(\lambda_{i+1}) \neq \emptyset$, choose any $\lambda \in e(\lambda_i) \cap e(\lambda_{i+1})$. Then,

$$B(\lambda_i) - k\theta^* \prec B(\lambda) \prec B(\lambda_i) + k\theta^*$$
$$B(\lambda_{i+1}) - k\theta^* \prec B(\lambda) \prec B(\lambda_{i+1}) + k\theta^*$$

Therefore, there must exist $k', k'' \in (-k, k)$ such that $B(\lambda_i) + k'\theta^* \sim B(\lambda) \sim B(\lambda_{i+1}) + k''\theta^*$, which implies $B(\lambda_{i+1}) + k''\theta^* \in \mathcal{A}'_i$. Then, this implies $B(\lambda_{i+1}) \in \mathcal{A}_{i+1}$. By induction, we conclude that $B(\lambda_i) \in \mathcal{A}_i$ for $i = 1, \ldots, n$, and also that $A \in \mathcal{A}'_n \subset \mathcal{A}_{n+1}$. Since $A \in \mathcal{I}$ was arbitrary, we have shown that $\mathcal{I} \subset \bigcup_i \mathcal{A}_i$. ■

We now examine $\mathcal{A}^c \setminus \bigcup_i \mathcal{A}_i$ more closely. First, note that if $C \in \mathcal{A}^c \setminus \bigcup_i \mathcal{A}_i$, then either $C \prec A$ for all $A \in \bigcup_i \mathcal{A}_i$ or $C \succ A$ for all $A \in \bigcup_i \mathcal{A}_i$. For suppose not. Then there exist $A, B \in \bigcup_i \mathcal{A}_i$ such that $A \succsim C \succsim B$. Now, there must be $\mathcal{A}_i, \mathcal{A}_{i'}$ such that $A \in \mathcal{A}_i$ and $B \in \mathcal{A}_{i'}$. Either $\mathcal{A}_{i'} \subset \mathcal{A}_i$ or $\mathcal{A}_i \subset \mathcal{A}_{i'}$. WLOG, suppose the former. Then, $A, B \in \mathcal{A}_i \subset \mathcal{A}'_i$, which, by Lemma 6, implies $C \in \mathcal{A}'_i \subset \bigcup_i \mathcal{A}_i$. This contradicts $C \in \mathcal{A}^c \setminus \bigcup_i \mathcal{A}_i$. Therefore, if we define

$$\begin{aligned} H &= \{A \in \mathcal{A}^c : A \succ B \ \forall B \in \bigcup_i \mathcal{A}_i\} \\ L &= \{A \in \mathcal{A}^c : A \prec B \ \forall B \in \bigcup_i \mathcal{A}_i\} \end{aligned}$$

then $H \cup L = \mathcal{A}^c \setminus \bigcup_i \mathcal{A}_i$. Note that it is possible that either H or L (or both) are empty. This would imply that $\bigcup_i \mathcal{A}_i = \mathcal{A}^c$, and $\hat{V} : \mathcal{A}^c \to \mathbb{R}$. We now prove that H and L each contain at most one indifference curve.

Lemma 9 *$A \sim A'$ for all $A, A' \in H$, and $B \sim B'$ for all $B, B' \in L$.*

Proof: We prove this property for H; the proof for L is similar. Suppose $A \succ A'$ for some $A, A' \in H$, and we will show there is a contradiction. Let $q = (\frac{1}{|Z|}, \ldots, \frac{1}{|Z|})$. By continuity, there must exist some $\lambda' \in (0,1)$ close enough to 1 that $\lambda' A + (1-\lambda')\{q\} \succ A'$. However, we can show that for all $\lambda \in [0,1)$, $\lambda A + (1-\lambda)\{q\} \in \mathcal{I}$. Consider an arbitrary $\theta \in \Theta$. By the choice of q, we know there exists a $k > 0$ such that $\{q\} + k\theta \in \mathcal{A}^c$. Let $k' = (1-\lambda)k$. Then,

$$\lambda A + (1-\lambda)\{q\} + k'\theta = \lambda A + (1-\lambda)(\{q\} + k\theta) \in \mathcal{A}^c.$$

Thus, for all $\lambda \in [0,1)$, $\lambda A + (1-\lambda)\{q\} \in \mathcal{I} \subset \bigcup_i \mathcal{A}_i$. Since $A' \in H$, this implies $\lambda' A + (1-\lambda')\{q\} \prec A'$, but this is a condradiction. ∎

By Lemma 9, we can extend $\hat{V}$ to $V : \mathcal{A}^c \to \bar{\mathbb{R}}$ if for all $A \in \mathcal{A}^c$ we define $V(A)$ as follows:

$$V(A) = \begin{cases} \sup \hat{V}(\bigcup_i \mathcal{A}_i) & \text{if } A \in H \\ \hat{V}(A) & \text{if } A \in \bigcup_i \mathcal{A}_i \\ \inf \hat{V}(\bigcup_i \mathcal{A}_i) & \text{if } A \in L \end{cases}$$

We now prove that V represents $\succsim$ on $\mathcal{A}^c$. This is accomplished by showing that $\sup \hat{V}(\bigcup_i \mathcal{A}_i) \in \hat{V}(\bigcup_i \mathcal{A}_i)$ implies $H = \emptyset$ and $\inf \hat{V}(\bigcup_i \mathcal{A}_i) \in \hat{V}(\bigcup_i \mathcal{A}_i)$ implies $L = \emptyset$. For suppose $\sup \hat{V}(\bigcup_i \mathcal{A}_i) \in \hat{V}(\bigcup_i \mathcal{A}_i)$. Then, there exists $A \in \bigcup_i \mathcal{A}_i$ such that $\hat{V}(A) = \sup \hat{V}(\bigcup_i \mathcal{A}_i)$. Thus $\hat{V}(A) \geq \hat{V}(B)$ for all $B \in \bigcup_i \mathcal{A}_i$, which implies $A \succsim B$ for all $B \in \bigcup_i \mathcal{A}_i$. Therefore, $\bigcup_i \mathcal{A}_i \subset W_A$, where we define $W_A \equiv \{B \in \mathcal{A}^c : A \succsim B\}$. Also, note that $L \subset W_A$. Therefore, $H \cup W_A = \mathcal{A}^c$. We also have $H \cap W_A = \emptyset$, and hence $H = \mathcal{A}^c \setminus W_A$. By the continuity of $\succsim$, both H and W_A are closed. Therefore, H is both open and closed, and since $W_A \neq \emptyset$, this is only possible if $H = \emptyset$. A similar argument can be used to show that $\inf \hat{V}(\bigcup_i \mathcal{A}_i) \in \hat{V}(\bigcup_i \mathcal{A}_i)$ implies $L = \emptyset$. We conclude that V represents $\succsim$ on $\mathcal{A}^c$.

Lemma 10 *V is continuous.*

Proof: We begin by proving that $co(\hat{V}(\bigcup_i \mathcal{A}_i)) = \hat{V}(\bigcup_i \mathcal{A}_i)$. This is accomplished by showing that $co(V_i(\mathcal{A}_i)) = V_i(\mathcal{A}_i)$ for all $i \geq 0$. Note that $V_0(\mathcal{A}_0) = v(\mathcal{S})$ is obviously convex.

Now suppose that $V_i(\mathcal{A}_i)$ is convex for $i \geq 0$, and we will show that $V_{i+1}(\mathcal{A}_{i+1})$ is also convex. Take any $\alpha \in co(V_{i+1}(\mathcal{A}_{i+1}))$, that is, any $\alpha \in \mathbb{R}$ such that there exist $s, t \in V_{i+1}(\mathcal{A}_{i+1})$ with $s \leq \alpha \leq t$. If $\alpha \in co(V_i(\mathcal{A}_i))$, then $\alpha \in V_i(\mathcal{A}_i) \subset V_{i+1}(\mathcal{A}_{i+1})$ by the induction assumption. Therefore, suppose $\alpha > \beta$ for all $\beta \in V_i(\mathcal{A}_i)$. (The proof for $\alpha < \beta$ for all $\beta \in V_i(\mathcal{A}_i)$ is similar.) Now, since $t \in V_{i+1}(\mathcal{A}_{i+1})$, there exists a $A \in \mathcal{A}_{i+1}$ such that $V_{i+1}(A) = t$. Since $A \in \mathcal{A}_{i+1}$, $A = B + k\theta^*$ for some $B \in \mathcal{A}'_i$, $k \in \mathbb{R}$. Note that $V'_i(\mathcal{A}'_i) = V_i(\mathcal{A}_i)$, which implies $V_{i+1}(B) = V'_i(B) < \alpha$. Thus $V_{i+1}(B) < \alpha \leq V_{i+1}(B + k\theta^*) = V_{i+1}(B) + kv \cdot \theta^*$, so there exists a $k' \in [0, k]$ such that $V_{i+1}(B + k'\theta^*) = V_{i+1}(B) + k'v \cdot \theta^* = \alpha$. Therefore, $\alpha \in V_{i+1}(\mathcal{A}_{i+1})$. By induction, $co(V_i(\mathcal{A}_i)) = V_i(\mathcal{A}_i)$ for all $i \geq 0$. This then implies that $co(\hat{V}(\bigcup_i \mathcal{A}_i)) = \hat{V}(\bigcup_i \mathcal{A}_i)$, for suppose $\alpha \in co(\hat{V}(\bigcup_i \mathcal{A}_i))$. Then, there exist $i, j \geq 0$ and $s \in V_i(\mathcal{A}_i)$, $t \in V_j(\mathcal{A}_j)$ such that $s \leq \alpha \leq t$. WLOG, suppose $j \geq i$. Then, $V_i(\mathcal{A}_i) \subset V_j(\mathcal{A}_j)$, so $s, t \in V_j(\mathcal{A}_j)$, which implies $\alpha \in V_j(\mathcal{A}_j) \subset \hat{V}(\bigcup_i \mathcal{A}_i)$.

We now use this result and the continuity of $\succsim$ to show that V is upper semicontinuous (u.s.c.), that is, $V^{-1}([a, +\infty])$ is closed for all $a \in \bar{\mathbb{R}}$. If $a \in V(\mathcal{A}^c)$, so that there exists a $A \in \mathcal{A}^c$ with $a = V(A)$, then $V^{-1}([a, +\infty]) = \{B \in \mathcal{A}^c : B \succsim A\}$, which is closed by the continuity of $\succsim$. However, we may not have $a \in V(\mathcal{A}^c)$. There are four cases to consider:

1. Suppose $a \leq \inf \hat{V}(\bigcup_i \mathcal{A}_i)$. Then, $V^{-1}([a, +\infty]) = \mathcal{A}^c$, which is closed.

2. Suppose $a > \sup \hat{V}(\bigcup_i \mathcal{A}_i)$. Then, $V^{-1}([a, +\infty]) = \emptyset$, which is closed.

3. Suppose $a = \sup \hat{V}(\bigcup_i \mathcal{A}_i)$. Then, $V^{-1}([a, +\infty]) = H$, which is closed (and also may be empty).

4. Suppose $\inf \hat{V}(\bigcup_i \mathcal{A}_i) < a < \sup \hat{V}(\bigcup_i \mathcal{A}_i)$. Then, there exist $s, t \in \hat{V}(\bigcup_i \mathcal{A}_i)$ such that $s < a < t$, which we showed implies $a \in \hat{V}(\bigcup_i \mathcal{A}_i) \subset V(\mathcal{A}^c)$. Therefore, as argued at the beginning of this paragraph, $V^{-1}([a, +\infty])$ is closed.

In proving that V is u.s.c., the necessary continuity assumption is that the upper contour sets, $\{A' \in \mathcal{A}^c : A' \succsim A\}$, are closed. By assuming that the lower contour sets, $\{A' \in \mathcal{A}^c : A' \precsim A\}$, are closed, a similar argument to that given above shows that V is lower semicontinuous (l.s.c.). Therefore, by assuming continuity of $\succsim$, we have that V is both u.s.c. and l.s.c., and hence continuous. ■

We now prove that V is not only θ^*-linear, but also Θ-linear.

Lemma 11 *V is Θ-linear.*

Proof: We first show that V is Θ-linear on $\mathcal{I}$. Then, we will use the continuity of V to show this implies V is Θ-linear on all of $\mathcal{A}^c$.

Suppose $A, A + \theta \in \mathcal{I}$ for some $\theta \in \Theta$. WLOG, suppose $A + \theta \succsim A$. (Otherwise, we can take $B = A + \theta$ and $\theta' = -\theta$.) Since $A, A + \theta \in \mathcal{I}$, there exist $k', k'' > 0$ such that

$A+k'\theta^*, A+\theta+k''\theta^* \in \mathcal{A}^c$. Let $k=\min\{k',k''\}$. Then, for all $\lambda\in[0,1]$, $A+\lambda\theta+k\theta^*\in\mathcal{A}^c$. This holds because

$$A+\lambda\theta+k\theta^* = \lambda(A+\theta+k\theta^*)+(1-\lambda)(A+k\theta^*),$$

and right side of the equation is a convex combination of two sets that are contained in $\mathcal{A}^c$. Since $A+k\theta^*\succ A$, by continuity there exists a $N\in\mathbb{N}$ large enough that $A+\frac{1}{N}\theta\precsim A+k\theta^*$. Since $A+\theta\succsim A$, Lemma 4 implies $A+\frac{1}{N}\theta\succsim A$, so we have $A\precsim A+\frac{1}{N}\theta\precsim A+k\theta^*$. Therefore, by continuity , there exists a $\bar{k}\in[0,k]$ such that $A+\frac{1}{N}\theta\sim A+\bar{k}\theta^*$. Now, for any $n\in\mathbb{N}$, $n<N$, TI-1 implies

$$A+\tfrac{n}{N}\theta+\tfrac{1}{N}\theta=(A+\tfrac{1}{N}\theta)+\tfrac{n}{N}\theta\sim(A+\bar{k}\theta^*)+\tfrac{n}{N}\theta=A+\tfrac{n}{N}\theta+\bar{k}\theta^*.$$

Therefore,

$$\begin{aligned}
V(A+\theta)-V(A) &= \sum_{n=0}^{N-1}\Big[V(A+\tfrac{n}{N}\theta+\tfrac{1}{N}\theta)-V(A+\tfrac{n}{N}\theta)\Big]\\
&= \sum_{n=0}^{N-1}\Big[V(A+\tfrac{n}{N}\theta+\bar{k}\theta^*)-V(A+\tfrac{n}{N}\theta)\Big]\\
&= \sum_{n=0}^{N-1} v\cdot\bar{k}\theta^*\\
&= N\bar{k}v\cdot\theta^*,
\end{aligned}$$

where the third equality follows from θ^*-linearity. We now need to show that $N\bar{k}v\cdot\theta^*=v\cdot\theta$. This is proved as follows:

$$\begin{aligned}
&A+\tfrac{1}{N}\theta\sim A+\bar{k}\theta^*=A+\tfrac{1}{N}\theta+(\bar{k}\theta^*-\tfrac{1}{N}\theta)\\
&\iff \{q\}\sim\{q\}+(\bar{k}\theta^*-\tfrac{1}{N}\theta) \qquad (q\in A+\tfrac{1}{N}\theta,\text{ by TI-2})\\
&\iff v\cdot(\bar{k}\theta^*-\tfrac{1}{N}\theta)=0 \qquad (\Theta\text{-linearity on } S)\\
&\iff N\bar{k}v\cdot\theta^*=v\cdot\theta.
\end{aligned}$$

Thus V is Θ-linear on $\mathcal{I}$.

Now, suppose $A, A+\theta\in\mathcal{A}^c$ for some $\theta\in\Theta$. Let $q=(\frac{1}{|Z|},\ldots,\frac{1}{|Z|})$. As shown in the proof of Lemma 9, for all $\lambda\in[0,1)$, $\lambda A+(1-\lambda)\{q\}\in\mathcal{I}$ and $\lambda(A+\theta)+(1-\lambda)\{q\}\in\mathcal{I}$. For all $n\in\mathbb{N}$, define $A_n\equiv(1-\frac{1}{n})A+\frac{1}{n}\{q\}$ and $\theta_n\equiv(1-\frac{1}{n})\theta$. Then $A_n\in\mathcal{I}$ for all $n\in\mathbb{N}$ and $A_n\to A$ as $n\to\infty$. Also, note that $A_n+\theta_n=(1-\frac{1}{n})(A+\theta)+\frac{1}{n}\{q\}$, so $A_n+\theta_n\in\mathcal{I}$ for all

$n \in \mathbb{N}$ and $A_n + \theta_n \to A + \theta$ as $n \to \infty$. Therefore,

$$\begin{aligned} V(A+\theta) - V(A) &= \lim_{n\to\infty} V(A_n + \theta_n) - \lim_{n\to\infty} V(A_n) \\ &= \lim_{n\to\infty} \big[V(A_n+\theta_n) - V(A_n)\big] \\ &= \lim_{n\to\infty} v \cdot \theta_n \\ &= v \cdot \lim_{n\to\infty} \theta_n \\ &= v \cdot \theta. \end{aligned}$$

Thus we see that V is Θ-linear on all of $\mathcal{A}^c$. ■

We now prove the last of the desired properties of V, namely convexity. This will also imply that V never attains $-\infty$, hence $V : \mathcal{A}^c \to \mathbb{R} \cup \{+\infty\}$ as claimed in Proposition 2. Before proceeding, define $\mathcal{C}$ to be the collection of all closed and bounded non-empty convex subsets of $\{p \in \mathbb{R}^Z : \sum_{z\in Z} p_z = 1\}$, endowed with the Hausdorff metric topology. Then $\mathcal{C}$ is complete since $\{p \in \mathbb{R}^Z : \sum_{z\in Z} p_z = 1\}$ is complete (see for instance Munkres, p279) and $\mathcal{A}^c \subset \mathcal{C}$.

Lemma 12 *V is convex.*

Proof: The argument given here is similar to that used in Lemma 20 of Maccheroni, Marinacci, and Rustichini (2004). We will first show that V is locally convex on $int(\mathcal{A}^c)$, where the interior is taken with respect to the Hausdorff topology on $\mathcal{C}$.[12] Also, it is easily verified that if we define $q = (\frac{1}{|Z|}, \ldots, \frac{1}{|Z|})$, then $\{q\} \in int(\mathcal{A}^c)$, so $int(\mathcal{A}^c) \neq \emptyset$.

Let $A_0 \in int(\mathcal{A}^c)$ be arbitrary. Then, there exists an $\epsilon > 0$ such that $B_\epsilon(A_0) \subset int(\mathcal{A}^c)$, where we define

$$B_\epsilon(A_0) = \{A \in \mathcal{C} : d_h(A, A_0) < \epsilon\}.$$

Note that $d_h(\cdot,\cdot)$ indicates the Hausdorff metric. Also, note that if $A \in \mathcal{C}$, then $A + \theta \in \mathcal{C}$ for all $\theta \in \Theta$ and $d_h(A, A+\theta) = \|\theta\|$, where $\|\cdot\|$ indicates the Euclidean norm. There exists $\theta \in \Theta$ such that $\|\theta\| < \epsilon$ and $A_0 + \theta \succ A_0$, and this implies that $A_0 + \theta \in B_\epsilon(A_0)$ and $v \cdot \theta > 0$. By continuity, there exists $\rho \in (0, \frac{1}{3})$ such that for all $A \in B_{\rho\epsilon}(A_0)$, $|V(A) - V(A_0)| < \frac{1}{3} v \cdot \theta$. Therefore, if $A, B \in B_{\rho\epsilon}(A_0)$, then

$$|V(A) - V(B)| \leq |V(A) - V(A_0)| + |V(A_0) - V(B)| < \tfrac{2}{3} v \cdot \theta.$$

[12] This step is necessary because $int(\mathcal{A}^c)$ with respect to the relative topology on $\mathcal{A}^c$ is $\mathcal{A}^c$ itself. That is, any topological space is itself open, so it is necessary for our purposes to consider $\mathcal{A}^c$ as a subset, not as a space.

Let $k \equiv \frac{V(A)-V(B)}{v\cdot\theta}$, which implies $|k| < \frac{2}{3}$. Then, we have

$$\begin{aligned} d_h(A_0, B + k\theta) &\leq d_h(A_0, B) + d_h(B, B + k\theta) \\ &< \rho\epsilon + \|k\theta\| \\ &< \tfrac{1}{3}\epsilon + \tfrac{2}{3}\epsilon = \epsilon, \end{aligned}$$

so $B + k\theta \in B_\epsilon(A_0) \subset \mathcal{A}^c$. Thus V is defined at $B + k\theta$. Note that $kv \cdot \theta = V(A) - V(B)$, so that $V(B + k\theta) = V(B) + kv \cdot \theta = V(A)$. Since $\succsim$ satisfies ACP, for any $\lambda \in (0, 1)$,

$$V(A) \geq V(\lambda A + (1-\lambda)(B + k\theta)).$$

Therefore,

$$\begin{aligned} V(A) &\geq V(\lambda A + (1-\lambda)B) + (1-\lambda)kv \cdot \theta \\ &= V(\lambda A + (1-\lambda)B) + (1-\lambda)(V(A) - V(B)), \end{aligned}$$

so we have

$$\lambda V(A) + (1-\lambda)V(B) \geq V(\lambda A + (1-\lambda)B).$$

Therefore, V is convex on $B_{\rho\epsilon}(A_0)$. Since the choice of $A_0 \in int(\mathcal{A}^c)$ was arbitrary, we see that each element of $int(\mathcal{A}^c)$ has a neighborhood on which V is convex. It is a standard result from convex analysis that this implies V is convex on $int(\mathcal{A}^c)$. Then, since $\mathcal{A}^c$ is convex and $int(\mathcal{A}^c) \neq \emptyset$, it is another standard result that $\overline{int(\mathcal{A}^c)} = \overline{\mathcal{A}^c} = \mathcal{A}^c$. Therefore, the continuity of V implies convexity on all of $\mathcal{A}^c$. ■

Lemma 13 *If $\succsim$ also satisfies strong continuity, then V is Lipschitz continuous.*

Proof: Let $\theta \in \Theta$ be as guaranteed by strong continuity and set $K \equiv v \cdot \theta$. We first show that for any $\epsilon > 0$ and $A, B \in \mathcal{A}^c$

$$A + \epsilon\theta, B + \epsilon\theta \in \mathcal{A}^c \;\&\; d_h(A, B) < \epsilon \;\Rightarrow\; |V(A) - V(B)| \leq K d_h(A, B). \tag{15}$$

To see this, let ϵ' be such that $d_h(A, B) < \epsilon' < \epsilon$. Since $\epsilon' \in (0, \epsilon)$, we have $A + \epsilon'\theta, B + \epsilon'\theta \in \mathcal{A}^c$, hence by strong continuity, $B + \epsilon'\theta \succ A$, i.e. $V(B + \epsilon'\theta) > V(A)$. This implies that $V(A) - V(B) < K\epsilon'$. A symmetric argument shows that $V(B) - V(A) < K\epsilon'$, implying that $|V(A) - V(B)| \leq K\epsilon'$. Since the latter holds for any ϵ' such that $d_h(A, B) < \epsilon' < \epsilon$, we conclude that $|V(A) - V(B)| \leq K d_h(A, B)$.

Next, we show that for any $\epsilon > 0$ and $A, B \in \mathcal{A}^c$

$$A + \epsilon\theta, B + \epsilon\theta \in \mathcal{A}^c \;\Rightarrow\; |V(A) - V(B)| \leq K d_h(A, B), \tag{16}$$

i.e. we do not actually need the requirement $d_h(A, B) < \epsilon$ to reach the conclusion in (15). To see this, let $C(\lambda) \equiv \lambda A + (1-\lambda)B$ for $\lambda \in [0,1]$. Note that $C(\lambda) + \epsilon\theta \in \mathcal{A}^c$, since

$$C(\lambda) + \epsilon\theta = \lambda(A + \epsilon\theta) + (1-\lambda)(B + \epsilon\theta).$$

By continuity of convex combinations, for each λ there exists an open (relative to $[0,1]$) interval $e(\lambda)$ such that $\lambda \in e(\lambda)$ and for all $\lambda' \in e(\lambda)$, $d_h(C(\lambda), C(\lambda')) < \frac{1}{2}\epsilon$. Thus $\{e(\lambda) : \lambda \in [0,1]\}$ is an open cover of $[0,1]$. Since $[0,1]$ is compact, there exists a finite subcover, $\{e(\lambda_1), \ldots, e(\lambda_k)\}$. Assume the λ_i's are arranged so that $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_k$. Define $\lambda_0 \equiv 0$ and $\lambda_{k+1} \equiv 1$. Note that if $i = 0, 1, \ldots, k$, then $d_h(C(\lambda_{i+1}), C(\lambda_i)) < \epsilon$, hence by (15)

$$|V(C(\lambda_{i+1})) - V(C(\lambda_i))| \leq K d_h(C(\lambda_{i+1}), C(\lambda_i)).$$

It is straightforward to verify that

$$d_h(C(\lambda_{i+1}), C(\lambda_i)) = (\lambda_{i+1} - \lambda_i) d_h(A, B).$$

Since $A = C(\lambda_{k+1})$ and $B = C(\lambda_0)$, by the triangular inequality and the above facts, we have

$$\begin{aligned} |V(A) - V(B)| &\leq \sum_{i=0}^{k} |V(C(\lambda_{i+1})) - V(C(\lambda_i))| \\ &\leq K \sum_{i=0}^{k} d_h(C(\lambda_{i+1}), C(\lambda_i)) \\ &= K \sum_{i=0}^{k} (\lambda_{i+1} - \lambda_i) d_h(A, B) \\ &= K d_h(A, B). \end{aligned}$$

To conclude the proof, let $A, B \in \mathcal{A}^c$ and define $q \equiv \left(\frac{1}{|Z|}, \ldots, \frac{1}{|Z|}\right)$. Then there exists a scalar $\kappa > 0$ such that $q + \kappa\theta \in \triangle(Z)$. For each integer n, let $A_n = \frac{1}{n}\{q\} + \left(1 - \frac{1}{n}\right) A$, $B_n = \frac{1}{n}\{q\} + \left(1 - \frac{1}{n}\right) B$, and $\epsilon_n = \frac{\kappa}{n}$. Then $A_n + \epsilon_n\theta, B_n + \epsilon_n\theta \in \mathcal{A}^c$. Hence by (16), $|V(A_n) - V(B_n)| \leq K d_h(A_n, B_n)$. Since $A_n \to A$, $B_n \to B$, and V is continuous, we have that $|V(A) - V(B)| \leq K d_h(A, B)$. ■

B.3 Application of the Fenchel-Moreau Duality

In this section we present three representations, each under alternative assumptions on V. Unlike in the RFCC representation, in these representations the set of measures $\mathcal{N}$ will not necessarily be minimal. We will derive a minimal representation in the next section.

In the first representation below, $\mathcal{N}$ need not be compact and supremum is taken over the set of measures instead of maximum. This will be used in the proof of the two subsequent representations and it is an interesting result in its own right. In the second and third representations, the set $\mathcal{N}$ can be chosen to be compact and the supremum can be replaced by the maximum. Additionally in the third representation, the measures in $\mathcal{N}$ can be guaranteed to be positive.

The remainder of this section is devoted to the proof of these results which are all contained in the following proposition. A function $V : \mathcal{A}^c \to \mathbb{R} \cup \{\infty\}$ is *monotone* if for all $A, B \in \mathcal{A}^c$ such that $A \subset B$, we have $V(A) \leq V(B)$.

Proposition 3 *Suppose that $V : \mathcal{A}^c \to \mathbb{R} \cup \{\infty\}$ satisfies 1–4 in Proposition 2. Then there exist a set of finite signed measures $\mathcal{N}$ and a lower semi-continuous function $c : \mathcal{N} \to \mathbb{R} \cup \{\infty\}$ such that*

1. *For all $A \in \mathcal{A}^c$*
$$V(A) = \sup_{\mu \in \mathcal{N}} \left[\int_{\mathcal{U}} \mu(du) \max_{q \in A} u \cdot q - c(\mu) \right]. \tag{17}$$
2. *$\mathcal{N}$ is consistent.*

If V is also Lipschitz continuous, then $\mathcal{N}$ can be chosen as a compact set and sup *can be replaced by* max. *If V is also Lipschitz continuous and monotone, then $\mathcal{N}$ can be chosen as a compact set of* positive *measures and* sup *can be replaced by* max.

In the remainder of this section assume that V satisfies 1–4 in Proposition 2. We will explicitly assume Lipschitz continuity and/or monotonicity of V at the end of this section, when we prove the stronger results. We will first extend V to $\mathcal{A}^c + \Theta$ by linearity. This will be useful in establishing consistency of the set of measures in the representation. Let $\mathcal{C}$ be defined as above, and we have $\mathcal{A}^c \subset \mathcal{A}^c + \Theta \subset \mathcal{C}$.

Lemma 14 *Let B_n be a Cauchy sequence in $\mathcal{A}^c + \Theta$. Then there is a subsequence B_{n_k}, sequences A_{n_k} in $\mathcal{A}^c$ and θ_{n_k} in Θ such that: $A_{n_k} \to A \in \mathcal{A}^c$, $\theta_{n_k} \to \theta \in \Theta$, and $B_{n_k} = A_{n_k} + \theta_{n_k} \to A + \theta$.*

Proof: The argument relies on $\mathcal{A}^c$ being compact and Θ being closed. Suppose that B_n is as in above, for any n choose $A_n \in \mathcal{A}^c$, $\theta_n \in \Theta$ such that $B_n = A_n + \theta_n$. Since $\mathcal{A}^c$ is a compact metric space, A_n has a convergent subsequence, w.l.o.g. A_n itself, converging to some $A \in \mathcal{A}^c$. Then:

$$\begin{aligned}
\|\theta_n - \theta_m\| &= d_h(\theta_n + A, \theta_m + A) \\
&\leq d_h(\theta_n + A, \theta_n + A_n) + d_h(\theta_n + A_n, \theta_m + A_m) \\
&\quad + d_h(\theta_m + A_m, \theta_m + A) \\
&= d_h(A, A_n) + d_h(B_n, B_m) + d_h(A_m, A) \to 0, \qquad \text{as } m, n \to \infty.
\end{aligned}$$

Hence θ_n is also convergent as a Cauchy sequence in the complete space Θ. Let $\theta \in \Theta$ denote its limit. We also have:

$$\begin{aligned} d_h(A_n + \theta_n, A + \theta) &\leq d_h(A_n + \theta_n, A + \theta_n) + d_h(A + \theta_n, A + \theta) \\ &= d_h(A_n, A) + \|\theta_n - \theta\| \to 0, \quad \text{as } n \to \infty, \end{aligned}$$

implying that $B_n = A_n + \theta_n \to A + \theta$.[13] ∎

Lemma 15 *$\mathcal{A}^c + \Theta$ is a convex and complete subspace of $\mathcal{C}$.*

Proof: The set $\mathcal{A}^c + \Theta$ is convex as the sum of two convex sets. To see that $\mathcal{A}^c + \Theta$ is a complete subspace of $\mathcal{C}$ it is enough to show that it is closed in $\mathcal{C}$, since $\mathcal{C}$ itself is complete. Let B_n be a sequence in $\mathcal{A}^c + \Theta$ that converges to some $B \in \mathcal{C}$. By Lemma 14, there is a subsequence B_{n_k} such that $B_{n_k} \to A + \theta$ for some $A \in \mathcal{A}^c$ and $\theta \in \Theta$, implying that $B = A + \theta \in \mathcal{A}^c + \Theta$. ∎

We next extend V to $\mathcal{A}^c + \Theta$, by setting $V(A + \theta) = V(A) + v \cdot \theta$ for any $A \in \mathcal{A}^c$ and $\theta \in \Theta$. We abuse notation and denote the extension of V to $\mathcal{A}^c + \Theta$ also by V. As shown in the next lemma, the extension is well-defined and preserves the properties of the original function.

Lemma 16 *The extension of V to $\mathcal{A}^c + \Theta$ is well-defined, convex, Θ-linear, and continuous. If the original function is Lipschitz continuous then the extension is Lipschitz continuous. If the original function is monotone then the extension is monotone.*

Proof: The extension is well-defined since if $A + \theta = A' + \theta'$ for $A, A' \in \mathcal{A}^c$ and $\theta, \theta' \in \Theta$, then

$$V(A) + v \cdot \theta = V(A' + [\theta' - \theta]) + v \cdot \theta = V(A') + v \cdot (\theta' - \theta) + v \cdot \theta = V(A') + v \cdot \theta'$$

by Θ-linearity of the original V on $\mathcal{A}^c$.

To see that the extension preserves convexity let $B = A + \theta$, $B' = A' + \theta'$ for some $A, A' \in \mathcal{A}^c$ and $\theta, \theta' \in \Theta$, then:

$$\begin{aligned} \lambda V(B) + (1-\lambda) V(B') &= \lambda[V(A) + v \cdot \theta] + (1-\lambda)[V(A') + v \cdot \theta'] \\ &= \lambda V(A) + (1-\lambda) V(A') + v \cdot (\lambda\theta + (1-\lambda)\theta') \\ &\geq V(\lambda A + (1-\lambda)A') + v \cdot (\lambda\theta + (1-\lambda)\theta') \\ &= V([\lambda A + (1-\lambda)A'] + [\lambda\theta + (1-\lambda)\theta']) \\ &= V(\lambda[A + \theta] + (1-\lambda)[A' + \theta']). \\ &= V(\lambda B + (1-\lambda)B'). \end{aligned}$$

[13]This is nothing but the continuity of $+ : \mathcal{A}^c \times \Theta \to \mathcal{C}$.

by Θ-linearity and convexity of the original V on $\mathcal{A}^c$.

To see that the extension preserves Θ-linearity, let $B = A + \theta$, $A \in \mathcal{A}^c$ and $\theta, \theta' \in \Theta$, then:

$$V(B + \theta') = V(A + [\theta + \theta']) = V(A) + v \cdot (\theta + \theta') = V(A) + v \cdot \theta + v \cdot \theta' = V(B) + v \cdot \theta'.$$

by Θ-linearity of the original V on $\mathcal{A}^c$.

Finally to see that the extension is continuous let $B_n \to B$ in $\mathcal{A}^c + \Theta$. By Lemma 14, there is a subsequence B_{n_k}, sequences A_{n_k} in $\mathcal{A}^c$ and θ_{n_k} in Θ such that: $A_{n_k} \to A \in \mathcal{A}^c$, $\theta_{n_k} \to \theta \in \Theta$, and $B_{n_k} = A_{n_k} + \theta_{n_k} \to A + \theta$. Hence $B = A + \theta$ and:

$$V(B_{n_k}) = V(A_{n_k}) + v \cdot \theta_{n_k} \to V(A) + v \cdot \theta = V(B)$$

by continuity of V on $\mathcal{A}^c$. This proves continuity of the extension.[14] ■

Next we follow a construction similar to the one in Dekel, Lipman, and Rustichini (2001) (henceforth DLR) to obtain from V, a function W whose domain is the set of support functions. For any $B \in \mathcal{C}$, the support function $\sigma_B : \mathcal{U} \to \mathbb{R}$ of B is defined by $\sigma_B(u) = \max_{p \in B} u \cdot p$. For a more complete introduction to support functions, see Rockafellar (1970) or Schneider (1993). Let $C(\mathcal{U})$ denote the set of continuous real-valued functions on $\mathcal{U}$ endowed with the supremum metric d_∞. Let $\Sigma = \{\sigma_B \in C(\mathcal{U}) : B \in \mathcal{A}^c + \Theta\}$. For any $\sigma \in \Sigma$ let

$$B_\sigma = \bigcap_{u \in \mathcal{U}} \left\{ p \in \triangle(Z) : u \cdot p = \sum_{z \in Z} u_z p_z \leq \sigma(u) \right\}.$$

Lemma 17 *1. For all $B \in \mathcal{A}^c + \Theta$ and $\sigma \in \Sigma$, $B_{(\sigma_B)} = B$ and $\sigma_{(B_\sigma)} = \sigma$. Hence σ is a bijection from $\mathcal{A}^c + \Theta$ to Σ.*

2. For all $B, B' \in \mathcal{A}^c + \Theta$, $\sigma_{\lambda B + (1-\lambda)B'} = \lambda \sigma_B + (1-\lambda)\sigma_{B'}$.

3. For all $B, B' \in \mathcal{A}^c + \Theta$, $d_h(B, B') = d_\infty(\sigma_B, \sigma_{B'})$.

Proof: These are standard results that can be found in Rockafellar (1972) or Schneider (1993). For instance in Schneider (1993), part 1 can be found in p39 (Theorem 1.7.1), part 2

[14] Let f be a function from a metric space X to a topological space Y. Then the following are equivalent: (i) f is continuous, (ii) for any sequence x_n in X, $x_n \to x$ implies $f(x_n) \to f(x)$, and (iii) for any sequence x_n in X, $x_n \to x$ implies that there is a subsequence x_{n_k} with $f(x_{n_k}) \to f(x)$. It is well-known that $(i) \iff (ii)$ and clearly $(ii) \Rightarrow (iii)$. In showing continuity of the extension in the above proof we used the fact that $(iii) \Rightarrow (i)$. The latter might not be as well known, so we show it here for completeness.

Suppose that (i) is not satisfied. Then f is not continuous at some $x \in X$. That is, there is a neighborhood $\mathcal{V}$ of $f(x)$ such that for any neighborhood $\mathcal{U}$ of x there is $x' \in \mathcal{U}$ with $f(x') \notin \mathcal{V}$. Hence for each n, we can choose x_n within $\frac{1}{n}$ distance of x such that $f(x_n) \notin \mathcal{V}$. Then $x_n \to x$, yet there is no subsequence x_{n_k} such that $f(x_{n_k}) \to f(x)$, so (iii) is not satisfied.

can be found in p37, and part 3 can be found in p53 (Theorem 1.8.11). ■

The set Σ is convex by part 2 of Lemma 17. The set $\mathcal{A}^c + \Theta$ is complete by Lemma 15, hence by parts 1 and 3 of Lemma 17, Σ is also complete. We conclude that Σ is closed as a complete subspace of the metric space $C(\mathcal{U})$.

Define the function $W : \Sigma \to \mathbb{R} \cup \{+\infty\}$ by $W(\sigma) = V(B_\sigma)$. We will say the function W is *monotone* if for all $\sigma, \sigma' \in \Sigma$ such that $\sigma \leq \sigma'$ we have $W(\sigma) \leq W(\sigma')$.

Lemma 18 *W is convex and continuous. If V is Lipschitz continuous with Lipschitz constant $K \geq 0$ then W is Lipschitz continuous with Lipschitz constant K. If V is monotone then W is monotone.*

Proof: To see that W is convex, let $B, B' \in \mathcal{A}^c + \Theta$, then:

$$W(\lambda\sigma_B + (1-\lambda)\sigma_{B'}) = W(\sigma_{\lambda B+(1-\lambda)B'}) = V(\lambda B + (1-\lambda)B')$$

$$\leq \lambda V(B) + (1-\lambda)V(B') = W(\sigma_B) + (1-\lambda)W(\sigma_{B'})$$

by parts 1 and 2 of Lemma 17 and convexity of V. Continuity of W follows from parts 1 and 3 of Lemma 17 and continuity of V. The function W inherits Lipschitz continuity of V with the same Lipschitz constant, by part 3 of Lemma 17. It inherits monotonicity of V because of the following fact which is easy to see from part 1 of Lemma 17: for all $B, B' \in \mathcal{A}^c + \Theta$, $B \subset B'$ iff $\sigma_B \leq \sigma_{B'}$. ■

We next extend W to $\tilde{W} : C(\mathcal{U}) \to \mathbb{R} \cup \{\infty\}$ by

$$\tilde{W}(f) = \begin{cases} W(f) & \text{if } f \in \Sigma \\ +\infty & \text{if } f \notin \Sigma \end{cases}$$

for all $f \in C(\mathcal{U})$. The extension is convex by convexity of the original function. It is lower-semi continuous since the original function is continuous on the closed domain Σ.

We denote the set of continuous linear functionals on $C(\mathcal{U})$ (the dual space of $C(\mathcal{U})$) by $C(\mathcal{U})^*$. It is well-known that $C(\mathcal{U})^*$ is the set of finite signed Borel measures on $\mathcal{U}$, where the duality is given by:

$$\langle f, \mu \rangle = \int_{\mathcal{U}} \mu(du) f(u)$$

for any $f \in C(\mathcal{U})$ and $\mu \in C(\mathcal{U})^*$.[15]

[15]Since $\mathcal{U}$ is a compact metric space, by the Riesz representation theorem (see e.g. Royden 1988, p357) each continuous linear function on $\mathcal{U}$ corresponds uniquely to a finite signed Baire measure on $\mathcal{U}$. Since $\mathcal{U}$ is a locally compact separable metric space, the Baire sets and the Borel sets of $\mathcal{U}$ coincide (Royden 1988, p332). Hence the set of Baire and Borel finite signed measures also coincide.

The polar function $\tilde{W}^*: C(\mathcal{U})^* \to \mathbb{R} \cup \{+\infty\}$ is defined by

$$\tilde{W}^*(\mu) = \sup_{f \in C(\mathcal{U})} \left[\langle f, \mu \rangle - \tilde{W}(f) \right].$$

The polar $\tilde{W}^*$ is convex and lower semi-continuous (see for example Proposition 3 in Ekeland and Turnbull 1983, p99). Note also that $\tilde{W}^*$ does not attain $-\infty$ since $\tilde{W}$ is not identically $+\infty$.

The bipolar function $\tilde{W}^{**}: C(\mathcal{U}) \to \mathbb{R} \cup \{-\infty, +\infty\}$ is defined by

$$\tilde{W}^{**}(f) = \sup_{\mu \in C(\mathcal{U})^*} \left[\langle f, \mu \rangle - \tilde{W}^*(\mu) \right].$$

Since $\tilde{W}$ is convex and lower-semi continuous, $\tilde{W} = \tilde{W}^{**}$ by the Fenchel-Moreau Theorem (see for example Proposition 1 in Ekeland and Turnbull 1983, p97). Note in particular that $\tilde{W}^*$ can not be identically $+\infty$ since otherwise $\tilde{W} = \tilde{W}^{**}$ would be identically $-\infty$.

Since $V(B) = \tilde{W}(\sigma_B)$, we have:

$$V(B) = \sup_{\mu \in C(\mathcal{U})^*} \left[\langle \sigma_B, \mu \rangle - \tilde{W}^*(\mu) \right] \tag{18}$$

for any $B \in \mathcal{A}^c + \Theta$.

By Proposition 2, the function $v : \triangle(Z) \to \mathbb{R}$ defined by $v(p) = V(\{p\})$ is in $\mathcal{U}$. Let $q = \left(\frac{1}{|Z|}, \ldots, \frac{1}{|Z|} \right)$, then $v(q) = 0$. For any $\mu \in C(\mathcal{U})^*$, define $u_\mu : \triangle(Z) \to \mathbb{R}$ by $u_\mu(p) = \langle \sigma_{\{p\}}, \mu \rangle$. It can be checked that u_μ is linear and $u_\mu(q) = 0$.

Lemma 19 *If $\mu \in C(\mathcal{U})^*$ is such that $u_\mu \neq v$, then $\tilde{W}^*(\mu) = +\infty$.*

Proof: Let $\mu \in C(\mathcal{U})^*$, $\lambda \in \mathbb{R}$ and set $\theta^\lambda = \lambda(u_\mu - v)$. Note that $\theta^\lambda \in \Theta$ since $\sum_{z \in Z} (u_\mu)_z = \sum_{z \in Z} v_z = 0$ from above. Pick any $p \in \triangle(Z)$. Then

$$V(\{p\}) + v \cdot \theta^\lambda = V(\{p + \theta^\lambda\}) \geq \langle \sigma_{\{p+\theta^l\}}, \mu \rangle - \tilde{W}^*(\mu) = u_\mu \cdot (p + \theta^\lambda) - \tilde{W}^*(\mu)$$

by $\Theta-$linearity of V and (18). This can further be simplified as:

$$-\lambda \|u_\mu - v\|^2 = -(u_\mu - v) \cdot \theta^\lambda \geq [u_\mu \cdot p - V(\{p\})] - \tilde{W}^*(\mu).$$

If $u_\mu \neq v$ then $\|u_\mu - v\| > 0$, so the first term tends to $-\infty$ as $\lambda \to \infty$. Since the term in the brackets on the right hand side is finite and independent of λ, this implies that $\tilde{W}^*(\mu) = +\infty$. ■

Let $\mathcal{N}_1$ denote the subset of $C(\mathcal{U})^*$ that consist of μ such that $u_\mu = v$ and let c denote the restriction of the polar $\tilde{W}^*$ to $\mathcal{N}_1$. Note that $\mathcal{N}_1$ is closed and c is lower semi-continuous.

Since $\mathcal{N}_1$ is a convex subset of $C(\mathcal{U})^*$, c is also convex. Moreover, by (18) and Lemma 19:

$$V(B) = \sup_{\mu \in \mathcal{N}_1} [\langle \sigma_B, \mu \rangle - c(\mu)] \tag{19}$$

for any $B \in \mathcal{A}^c + \Theta$. Since $\mathcal{A}^c \subset \mathcal{A}^c + \Theta$, this completes the proof of the first part of Proposition 3.

Lemma 20 *Let $B \in \mathcal{A}^c + \Theta$, if $\mu \in \partial(W(\sigma_B))$, then $\mu \in \mathcal{N}_1$ and μ is a maximizer of (19).*

Proof: If $\mu \in \partial(W(\sigma_B))$, then by the definition of the subdifferential, for all $\sigma \in \Sigma$ we have

$$\langle \sigma, \mu \rangle - W(\sigma) \leq \langle \sigma_B, \mu \rangle - W(\sigma_B).$$

Hence by the definition of the polar of the extension, $\tilde{W}^*(\mu) = \langle \sigma_B, \mu \rangle - W(\sigma_B)$. Since $\tilde{W}^*(\mu)$ is finite, by Lemma 19 $\mu \in \mathcal{N}_1$. Hence

$$V(B) \equiv W(\sigma_B) = \langle \sigma_B, \mu \rangle - \tilde{W}^*(\sigma_B)$$

so μ is a maximizer of (19). ■

We next prove the second part of Proposition 3.

Lemma 21 *Suppose that W is Lipschitz continuous with Lipschitz constant $K \geq 0$. Define*

$$\mathcal{N}_2 = \{\mu \in \mathcal{N}_1 \mid \|\mu\| \leq K\}.$$

Then $\mathcal{N}_2$ is a compact set of measures, equation (19) holds if we replace $\mathcal{N}_1$ by $\mathcal{N}_2$ and supremum by maximum.

Proof: By Alaoglu's Theorem the closed ball of the norm dual $B_K = \{\mu \in C(\mathcal{U})^* \mid \|\mu\| \leq K\}$ is weak*-compact. Hence $\mathcal{N}_2$ is weak*-closed as the intersection of weak*-closed sets $\mathcal{N}_1$ and B_K, and it is norm-bounded as a subset of the norm bounded set B_K. By Theorem 6.25 in Aliprantis and Border (1999, p250) $\mathcal{N}_2$ is weak*-compact as a weak*-closed and norm bounded set in the norm dual.

To finish the proof let $B \in \mathcal{A}^c + \Theta$. By Lemma 1 there exists $\mu \in \partial W(\sigma_B)$ such that $\|\mu\| \leq K$. By Lemma 20, $\mu \in \mathcal{N}_2$ and μ is a maximizer of (19). ■

We next prove a separation result which to a large extent mimicks Lemma 1. Unlike Lemma 1, the following result only applies to the Banach space $C(\mathcal{U})$, but it gives a stronger conclusion when the function is monotone. We will use it to prove the last claim in Proposition 3, the same way we used Lemma 1 to prove the second claim in Proposition 3. Let $C_+(\mathcal{U}) \equiv \{f \in C(\mathcal{U}) \mid f \geq 0\}$.

Lemma 22 *Assume that, in addition its convexity, $W : \Sigma \to \mathbb{R}$ is also Lipschitz continuous and monotone. If $K \geq 0$ is the Lipschitz constant of W, then for all $\sigma \in \Sigma$, there exists a positive measure $\mu \in \partial W(\sigma)$ with $\|\mu\| \leq K$.*

Proof: The epigraph of W is defined as:

$$epi(W) = \{(f,t) \in C(\mathcal{U}) \times \mathbb{R} \ : \ t \geq W(f), f \in \Sigma\}.$$

Note that $epi(W) \subset C(\mathcal{U}) \times \mathbb{R}$ is a convex set because W is convex with the convex domain Σ. Now, define

$$H = \{(f,t) \in C(\mathcal{U}) \times \mathbb{R} \ : \ t < -K\|f\|\}.$$

It is easily seen that H is nonempty and convex. Also, since $\|\cdot\|$ is necessarily continuous, H is open in the product topology. Define

$$I = H + C_+(\mathcal{U}) \times \{0\}.$$

Then $I \subset C(\mathcal{U}) \times \mathbb{R}$ is convex as the some of two convex sets, and it has non-empty interior since it contains the nonempty open set H.

Let $\sigma \in \Sigma$ be arbitrary. Let $I(\sigma) = (\sigma, W(\sigma)) + I$. We claim that $epi(W) \cap I(\sigma) = \emptyset$. To see this, note first that

$$I(\sigma) \ = \ \{(\sigma + f + g, W(\sigma) + t) \ : \ t < -K\|f\|, t \in \mathbb{R}, f \in C(\mathcal{U}), \text{ and } g \in C_+(\mathcal{U})\}.$$

Suppose for a contradiction that $(\sigma', s) \in epi(W) \cap I(\sigma)$. Then $\sigma' \in \Sigma$, $s \geq W(\sigma')$, and there exist $f \in C(\mathcal{U})$, $g \in C_+(\mathcal{U})$ such that $\sigma' = \sigma + f + g$, and $s - W(\sigma) < -K\|f\|$. By Lipschitz continuity, we have $W(\sigma') \geq W(\sigma) - K\|\sigma' - \sigma\|$. Therefore, $s \geq W(\sigma) - K\|\sigma' - \sigma\|$, which implies $(\sigma', s) \notin H(\sigma)$.

We showed that $I(\sigma)$ and $epi(W)$ are disjoint convex sets and $I(\sigma)$ has nonempty interior. Therefore, a version of the Separating Hyperplane Theorem implies there exists a nonzero continuous linear functional $(\mu, \lambda) \in C(\mathcal{U})^* \times \mathbb{R}$ that separates $I(\sigma)$ and $epi(W)$.[16] That is, there exists a scalar δ such that

$$\langle f, \mu \rangle + \lambda t \leq \delta \quad \text{if} \quad (f,t) \in epi(W) \tag{20}$$

and

$$\langle f, \mu \rangle + \lambda t \geq \delta \quad \text{if} \quad (f,t) \in I(\sigma). \tag{21}$$

Clearly, we cannot have $\lambda > 0$. Also, if $\lambda = 0$, then Equation (21) implies $\mu = 0$. This would contradict (μ, λ) being a nonzero functional. Therefore, $\lambda < 0$. Without loss of generality, we can take $\lambda = -1$, for otherwise we could renormalize (μ, λ) by dividing by $-\lambda$.

[16] See Aliprantis and Border (1999, Theorem 5.50) or Luenberger (1969, p. 133).

Since $(\sigma, W(\sigma)) \in epi(W)$, we have $\langle \sigma, \mu \rangle - W(\sigma) \leq \delta$. For all $t > 0$, we have $(\sigma, W(\sigma) - t) \in I(\delta)$, which implies $\langle \sigma, \mu \rangle - W(\sigma) + t \geq \delta$. Therefore, $\langle \sigma, \mu \rangle - W(\sigma) = \delta$, and thus for all $\sigma' \in \Sigma$,

$$\langle \sigma', \mu \rangle - W(\sigma') \leq \delta = \langle \sigma, \mu \rangle - W(\sigma).$$

Equivalently, we can write $W(\sigma') - W(\sigma) \geq \langle \sigma' - \sigma, \mu \rangle$. Thus, $\mu \in \partial W(\sigma)$.

To see that $\|\mu\| \leq K$, suppose to the contrary. Then, there exists $f \in C(\mathcal{U})$ such that $\langle f, \mu \rangle < -K\|f\|$, and hence there also exists $\epsilon > 0$ such that $\langle f, \mu \rangle + \epsilon < -K\|f\|$. Therefore,

$$\langle f + \sigma, \mu \rangle - W(\sigma) + K\|f\| + \epsilon < \langle \sigma, \mu \rangle - W(\sigma) = \delta,$$

which, by Equation (21), implies $(f + \sigma, W(\sigma) - K\|f\| - \epsilon) \notin I(\sigma)$. However, this contradicts the definition of $I(\sigma)$. Thus $\|\mu\| \leq K$.

It only remains to show that μ is a positive measure. Let $f \in C_+(\mathcal{U})$. Then for any $\epsilon > 0$, $(\sigma + f, W(\sigma) - \epsilon) \in I(\sigma)$. By equation (21)

$$\langle \sigma + f, \mu \rangle - W(\sigma) + \epsilon \geq \delta = \langle \sigma, \mu \rangle - W(\sigma),$$

hence $\langle f, \mu \rangle \geq -\epsilon$. Since the latter holds for all $\epsilon > 0$ and $f \in C_+(\mathcal{U})$, we have that $\langle f, \mu \rangle \geq 0$ for all $f \in C_+(\mathcal{U})$. Therefore μ is a positive measure. ■

Lemma 23 *Suppose that W is Lipschitz continuous and monotone with Lipschitz constant $K \geq 0$. Define $\mathcal{N}_3$ as the set of positive measures in $\mathcal{N}_2$. Then $\mathcal{N}_3$ is a compact set of measures, equation (19) holds if we replace $\mathcal{N}_1$ by $\mathcal{N}_3$ and supremum by maximum.*

Proof: The set of positive measures is closed and by Lemma 21 $\mathcal{N}_2$ is compact. Hence their intersection $\mathcal{N}_3$ is also compact. Let $B \in \mathcal{A}^c + \Theta$. By Lemma 22 there exists a positive measure $\mu \in \partial W(\sigma_B)$ such that $\|\mu\| \leq K$. By Lemma 20, $\mu \in \mathcal{N}_3$ and μ is a maximizer of (19). ■

B.4 Derivation of a Minimal $\mathcal{M}$

In this section we show that Zorn's Lemma guarantees the existence of a minimal set of measures.

Proposition 4 *Suppose that $V : \mathcal{A}^c \to \mathbb{R} \cup \{\infty\}$, $\mathcal{N}$ is a compact and consistent set of signed measures, and $c : \mathcal{N} \to \mathbb{R} \cup \{\infty\}$ is a lower semi-continuous function such that for all $A \in \mathcal{A}^c$*

$$V(A) = \max_{\mu \in \mathcal{N}} [\langle \sigma_A, \mu \rangle - c(\mu)]. \tag{22}$$

Then there exists a compact subset $\mathcal{M}$ of $\mathcal{N}$ such that: (i) equation (22) continues to hold if we replace $\mathcal{N}$ by $\mathcal{M}$ and (ii) equation (22) does not hold if we replace $\mathcal{N}$ by any proper closed subset $\mathcal{M}'$ of $\mathcal{M}$.

Proof: Let ξ denote the collection of all nonempty compact subsets $\mathcal{K}$ of $\mathcal{N}$ for which equation (22) holds if we replace $\mathcal{N}$ by $\mathcal{K}$. We will show that ξ has a minimal element with respect to set inclusion, which will conclude the proof. Let ζ be a subset of ξ whose elements are linearly ordered with respect to set inclusion. If we can show that there is an $\hat{\mathcal{K}} \in \xi$ such that $\hat{\mathcal{K}} \subset \mathcal{K}$ for all $\mathcal{K} \in \zeta$, then we can invoke Zorn's lemma to conclude that ξ has a minimal element.

Let $\hat{\mathcal{K}} \equiv \cap\{\mathcal{K} \,|\, \mathcal{K} \in \zeta\}$. Then $\hat{\mathcal{K}}$ is nonempty, since $\mathcal{N}$ is compact and ζ consists of closed subsets of $\mathcal{N}$ which satisfy the finite intersection property (see Theorem 2.28 in Aliprantis and Border 1999, p38). We also have that $\hat{\mathcal{K}}$ is compact since it is closed as the intersection of closed sets and norm bounded as the subset of the norm bounded set $\mathcal{N}$ (see Theorem 6.25 in Aliprantis and Border, 1999, p250). To conclude that $\hat{\mathcal{K}} \in \xi$, we will show that equation (22) holds if we replace $\mathcal{N}$ by $\hat{\mathcal{K}}$.

Fix $A \in \mathcal{A}^c$. For each $\mathcal{K} \in \zeta$ choose $\mu(\mathcal{K}) \in \mathcal{K}$ that maximizes (22) when we replace $\mathcal{N}$ by $\mathcal{K}$. Then

$$c(\mu(\mathcal{K})) = \langle \sigma_A, \mu(\mathcal{K}) \rangle - V(A).$$

Moreover $(\mu(\mathcal{K}))_{\mathcal{K}\in\zeta}$ forms a net in $\mathcal{N}$. Since $\mathcal{N}$ is compact this net must have a convergent subnet $(\mu(\mathcal{K}))_{\mathcal{K}\in\zeta'}$ (see Theorem 2.28 in Aliprantis and Border 1999, p38) with a limit μ. For all $\mathcal{K} \in \zeta$, μ is a limit point of $\{\mu(\mathcal{K}') \,|\, \mathcal{K}' \in \zeta, \mathcal{K}' \subset \mathcal{K}\}$, hence by compactness of $\mathcal{K}$, $\mu \in \mathcal{K}$. Therefore $\mu \in \hat{\mathcal{K}}$.

By lower semi continuity of c (see Theorem 2.39 in Aliprantis and Border 1999, p43)

$$c(\mu) \leq \liminf_{\mathcal{K}} c(\mu(\mathcal{K}))$$

where the above limit is taken on the subnet $(\mu(\mathcal{K}))_{\mathcal{K}\in\zeta'}$ converging to μ. The above two equations imply that

$$V(A) \leq \langle \sigma_A, \mu \rangle - c(\mu).$$

Since μ was shown to be in $\hat{\mathcal{K}}$, the last equation implies that (22) holds if we replace $\mathcal{N}$ by $\hat{\mathcal{K}}$. ■

References

[1] Aliprantis, C., and K. Border (1999): *Infinite Dimensional Analysis*. Berlin, Germany: Springer-Verlag.

[2] Dekel, E., B. Lipman, and A. Rustichini (2001): "Representing Preferences with a Unique Subjective State Space," *Econometrica*, 69, 891–934.

[3] Dekel, E., B. Lipman, and A. Rustichini (2005): "Temptation-Driven Preferences," Working Paper, Boston University.

[4] Dekel, E., B. Lipman, A. Rustichini, and T. Sarver (2005): "Representing Preferences with a Unique Subjective State Space: Corrigendum," mimeo, Boston University.

[5] Ekeland, I., and T. Turnbull (1983): *Infinite-dimensional Optimization and Convexity*, Chicago: The University of Chicago Press.

[6] Ergin, H. (2003), "Costly Contemplation," mimeo, MIT.

[7] Gul, F., and W. Pesendorfer (2001): "Temptation and Self-Control," *Econometrica*, 69, 1403–1435.

[8] Hörmander, L. (1954): "Sur la Fonction d'Appui des Ensembles Convexes dans un Espace Localement Convexe," *Arkiv för Matematik*, 3, 181–186.

[9] Kreps, D. (1979): "A Representation Theorem for Preference for Flexibility," *Econometrica*, 47, 565–578.

[10] Luenberger, D. G. (1969), *Optimization by Vector Space Methods*, John Wiley and Sons, Inc., New York.

[11] Mas-Colell, A. (1986): "The Price Equilibrium Existence Problem in Topological Vector Lattices," *Econometrica*, 54, 1039–1054.

[12] Phelps, R. R. (1992): *Convex Functions, Monotone Operators, and Differentiability*, Berlin, Germany: Springer-Verlag.

[13] Rockafellar, R. T. (1970): *Convex Analysis*. Princeton, NJ: Princeton University Press.

[14] Royden, H. L. (1988): *Real Analysis*. Englewood Cliffs, NJ: Prentice Hall.

[15] Sarver, T. (2005), "Anticipating Regret: Why Fewer Options May Be Better," mimeo, Boston University.

[16] Schneider, R. (1993): *Convex Bodies: The Brunn-Minkowski Theory*. Cambridge: Cambridge University Press.